DOCUMENT RESUME

ED 025 781 CG 002 334

By-Goldstein, Harris K.

Identifying and Maximizing Research Learning Potential for Social Work Students. Tulane Studies in Social Welfare.

Tulane Univ. of Louisiana, New Orleans. School of Social Work.
Pub Date Oct 67
Note-100p.
EDRS Price MF-$0.50 HC-$5.10
Descriptors-*Graduate Students, *Graduate Study, *Pilot Projects, *Research Skills, *Social Work, Social Workers, Student Research

A pilot project sought to learn how, at the time of admission, graduate schools of social work can begin to identify levels of potential and flair for, and interest in, doing research. Data of 353 students were obtained before these students had any classroom training in research. Data on 263 students were obtained after the course was completed. The training plan exploited a number of naturally occurring differences among the five schools that influenced student learning. These included such factors as the relative emphasis on various aspects of content, differences in methods of teaching, and in learning experiences, variations in methods of evaluating learning achieved, the time and place of research in the curriculum, and teacher characteristics. Data on the objectives, methods of teaching research, course content, and other details were obtained. Student characteristics and biographical data were also obtained. An instrument, the "Measurement of Attitudes and Research Knowledge" (MARK), assessed student attitude to, and knowledge of, research prior to the course and subsequent to its completion. Students also evaluated the teaching. The study found that various levels of potential for learning research can be identified, and learning achieved by students can be predicted. (Author/EK)

ED025781

U.S. DEPARTMENT OF HEALTH, EDUCATION & WELFARE
OFFICE OF EDUCATION

THIS DOCUMENT HAS BEEN REPRODUCED EXACTLY AS RECEIVED FROM THE PERSON OR ORGANIZATION ORIGINATING IT. POINTS OF VIEW OR OPINIONS STATED DO NOT NECESSARILY REPRESENT OFFICIAL OFFICE OF EDUCATION POSITION OR POLICY.

TULANE STUDIES

in

SOCIAL WELFARE

Identifying and Maximizing Research Learning Potential for Social Work Students

CG 002 334

School of Social Work

Tulane University

Vol. IX, New Orleans, La., 1967

IDENTIFYING AND MAXIMIZING RESEARCH LEARNING POTENTIAL

FOR

SOCIAL WORK STUDENTS

Harris K. Goldstein, D.S.W.

Professor of Social Work

TULANE UNIVERSITY

October, 1967

TABLE OF CONTENTS

Chapter		Page
ACKNOWLEDGEMENTS		vii
I.	INTRODUCTION	1
II.	STUDIES OF RESEARCH TEACHING AND SCIENTIFIC PRODUCTIVITY, AND THEIR RELATION TO STUDY METHODOLOGY.	7
	Previous Concerns With Research Teaching	7
	Theoretical and Empirical Bases for Predicting Success in Research.	11
	Identifying Research Potential in This Training Plan	14
III.	METHODOLOGY OF THE TRAINING PLAN	17
	Method of Obtaining Data on Students Prior to Their Taking Research Course	17
	The I.D. Questionnaire	18
	The Interest Test.	19
	The "Before" MARK.	20
	Method of Obtaining Data on Students' Subsequent to Their Completing the Research Course	21
	The "After" MARK	22
	The Special Test Questionnaire	23
	Course Evaluation.	24
	Course Grades.	25
	Methods of Obtaining Data on Teaching Variables.	26
	School Catalogs.	26
	Written Material from Teachers	26
	Correspondence and Interviews with Assistant Project Directors.	26
IV.	STUDENT CHARACTERISTICS PRIOR TO TRAINING.	27
	Comparison of This Sample with Students in Other Schools of Social Work	27
	Typical Student Profile.	28
	I.D. Variables	30
	Parents' Income, Occupation, and Education	30
	Cultural Background.	30
	Educational Background	31
	Background in Mathematics and Abstract Reasoning	31
	Spare Time Activities and Interests.	32
	Background in Research	33

TABLE OF CONTENTS (Cont'd.)

Chapter		Page
V.	THE TEACHING VARIABLES IN THE TRAINING PLAN.	35
	Differences and Similarities in Objectives, Methods, and Content.	35
	Objectives	37
	Methods.	38
	Content.	40
	Teacher Characteristics.	41
	Miscellaneous Variables.	41
	Experimentation Introduced as Part of the Pilot Training Plan.	42
VI.	EVALUATION OF INSTRUMENTATION AND DATA COLLECTION.	43
	The MARK	43
	Reliability--"Before" and "After" MARK	46
	Item Difficulty and Heterogeneity "Before" MARK. . . .	47
	Discrimination	48
	Validity	49
	Item Difficulty and Heterogeneity--"After" MARK. . . .	51
	The I.D. Questionnaire and Interest Test	52
	The Special Test Questions	53
	The Course Evaluation.	54
	Research Course Grades	54
VII.	THE OUTCOME AND EVALUATION OF THE PROJECT.	56
	Identification of Levels of Student Potential for Learning Research.	57
	Identification and Prediction by I.D. and Interest Data.	57
	Identification of Potential Levels by "Before" MARK Scores.	60
	Prediction of Learning by "Before" MARK Scores	61
	Prediction of "After" from "Before" Scores	61
	Prediction of Grades from "Before" Scores.	61
	Prediction of Special Test Questions Scores from "Before" Scores.	62
	The Relation of the Teaching Variables to Various Levels of Potential.	63
	Similarity of the "Before" and "After" Sample.	64
	Differences Between "Before" and "After" MARK Scores	65
	Relation of Increase in MARK Score to Students' Rank on "Before" MARK.	67
	Relationship Between Changes in MARK Scores and Teaching Variables at the Five Schools	69

TABLE OF CONTENTS (Cont'd.)

Chapter		Page
VII.	THE OUTCOME AND EVALUATION OF THE PROJECT (Cont'd.)	
	Teaching Methods.	70
	Learning Experiences.	70
	Evaluating Learning Achieved.	71
	Content	71
	Teacher Characteristics	71
	Miscellaneous Variables	71
	Relationship Between Changes in MARK Scores and Teaching Variables in the Fourteen Sections.	72
	Relation Between Change in MARK Scores and Special Test Questions.	73
	Relation of Special Test Questions Scores and Teaching Variables by Schools.	74
	Students' Responses to the Course Evaluation.	75
	Students' Assessment of the Knowledge They Gained . . .	75
	Usefulness of Students' Learning to Their Practice. . .	75
	Students' Interest in the Course.	75
	Students' Reaction to Teaching Methods.	75
	Students' Characterization of Course Content.	76
	Relationship Between Change in MARK Scores and Students' Evaluation of Course.	78
	Relationship Between Course Grades and Course Evaluations	78
	Student Assessment of the Knowledge They Gained . . .	78
	Usefulness to Practice and Interest in Research . . .	79
	Reaction to Teaching Methods.	79
	Students' Characterization of Course Content.	79
	Relationship of Students' Evaluation of Learning and Teaching Variables by Schools	81
	Student Assessment of Knowledge They Gained	81
	Usefulness to Practice and Interest in Research . . .	81
	Reaction to Teaching Methods.	81
	Students' Characterization of Course Content.	82
VIII.	DISCUSSION OF FINDINGS AND SUGGESTIONS FOR THE FUTURE . . .	83
	Three Kinds of Student Potential for Learning Research. .	83
	A Suggested Curriculum to Maximize Student Potential. . .	87
	Suggestions for Future Training Plans	90

LIST OF TABLES

Number		Page
1	MEAN SCORES ON "BEFORE" MARK TEST BY SCHOOLS.	34
2	MEDIAN SCORES ON "BEFORE" MARK TEST BY SCHOOLS.	34
3	COURSE OBJECTIVES	38
4	METHODS OF TEACHING	39
5	METHODS OF EVALUATING LEARNING.	39
6	LEARNING EXPERIENCES.	40
7	CONTENT OF RESEARCH COURSES	40
8	TEACHER CHARACTERISTICS	41
9	MISCELLANEOUS VARIABLES	41
10	RELIABILITY OF "BEFORE" AND "AFTER" MARK SCORES	46
11	LEVEL OF ITEM DIFFICULTY--"BEFORE" MARK	48
12	ITEM DISCRIMINATION--"BEFORE" MARK.	49
13	VALIDITY OF "AFTER" MARK SCORES AS SHOWN BY CORRELATION WITH RESEARCH COURSE GRADES	50
14	LEVEL OF ITEM DIFFICULTY--"AFTER" MARK.	51
15	ITEM DISCRIMINATION--"AFTER" MARK	52
16	CORRELATIONS OF INTEREST AND I.D. ITEMS WITH COURSE GRADES, "AFTER" MARK SCORES, AND SPECIAL TEST QUESTIONS SCORES.	59
17	RELATION BETWEEN STUDENTS' RANK ON "BEFORE" OR "AFTER" MARK.	60
18	PREDICTIVE POWERS OF "BEFORE" MARK SCORES WITH THE "AFTER" MARK SCORES AS A CRITERION.	61
19	PREDICTIVE POWER OF "BEFORE" MARK SCORES WITH COURSE GRADES AS A CRITERION	62

LIST OF TABLES (Cont'd.)

NUMBER		Page
20	PREDICTIVE POWER OF THE "BEFORE" MARK WITH SPECIAL TEST QUESTIONS AS A CRITERION.	63
21	INCREASE IN TOTAL MARK SCORES AFTER STUDENTS HAD TAKEN THE RESEARCH COURSE.	65
22	INCREASE IN MEAN ATTITUDE AND KNOWLEDGE SCORES AFTER STUDENTS HAD TAKEN THE RESEARCH COURSE	66
23	INCREASE IN MARK KNOWLEDGE SCORES BY STUDENTS' RANK. . .	67
24	INCREASE IN MARK SCORES--STUDENTS SECTIONED BY RANK. . .	68
25	SCHOOL TO SCHOOL COMPARISONS ON STUDENT CHARACTERIZATION OF COURSE CONTENT.	77
26	CHARACTERIZATION OF COURSE CONTENT BY STUDENT LEVEL. . .	80

ACKNOWLEDGEMENTS

In the pilot training project reported here, like most efforts aimed at studying the training of social work students, there is much indebtedness to many people. Particular mention should be made of the ten research teachers at Maryland, Wisconsin, Denver, and Tennessee, whose classes (together with the writer's) provided data for this project. Mary H. Block of the University of Tennessee, Harris Chaiklin of the University of Maryland, James Kern of the University of Denver, and Morton Perlmutter of the University of Wisconsin, at Madison, were of much help in their capacity as Assistant Directors of the project.

The research apprentices who worked with the writer for the last four years provided considerable assistance in the preliminary work on the questions and instruments used to evaluate this project. Without their previous help, it would not have been possible to complete this study in the two years allocated to it. This group included Douglas Radabaugh, Julanne Haspel, Faye Weckel, Eugene (Jake) Hadley, Mary Jane Winn, and Louisa Dodkin. Terry Gorman Salassi and Ron Burris were not only included in the group of apprentices who did previous work with the writer, but played a major role in helping in the analysis of the evaluative material.

The writer is grateful to personnel in the Tulane Computer Center, and particularly to Mrs. Loato Wissner, the Administrative Assistant. Without the assistance of the Center in programming and analysis of data, the numerous calculations required for the project would have consumed much more time and money. Mrs. Olive Dudenhefer typed the report.

The Deans of the participating schools, Walter L. Kindelsperger, Verl S. Lewis, Martin B. Loeb, Sue Spencer, and Emil M. Sunley, assisted this project not only by allocating time of faculty and clerical staff, but supported it also in many intangible ways. The National Institute of Mental Health supported this study with their pilot training grant No. 1TI-MH-10710-01.

The writer alone is responsible for any deficiencies in design or analysis, and the clarity of the written report.

CHAPTER I

INTRODUCTION

Development of this pilot training project was stimulated by observation of the growing demand in social work for persons trained to do research and the relatively limited special efforts being made to meet this need. The writer knows of no formal study of the relative number of positions for whom social work researchers are being sought, nor the relative number of social work researchers looking for jobs. There appears to be general agreement, however, that there is a serious shortage.

That social workers enter the field largely because they are interested in working with people and want to provide service to them, rather than because of their interest in developing knowledge, is an assumption which few social workers will dispute. At least one formal study supports this assumption.[1]

The official recruitment efforts of the profession also stress the opportunities for giving service rather than the opportunities for developing knowledge.[2] This emphasis on the helping possibilities in

---

[1]Arnulf Pins, "Who Chooses Social Work, When and Why," p. 87, shows that 62 per cent of the persons who entered social work in 1960 entered because they considered they would be making a contribution to an individual or society, or because they liked to work with people. None gave answers that showed they wanted to learn how to contribute to developing knowledge, but this may have been due to the form of the questions.

[2]See "Summer of Decision," by the Council of Social Work Education, a 35 mm film, 1959.

the field is certainly valid for a profession whose chief reason for being is to provide assistance for people in distress, providing it does not have as a corollary any nonemphasis on the development of knowledge on which competent professional performance is based.

The pilot training plan described in this report assumed that the focus of recruitment efforts of the profession would remain the same in the near future; that is, it expected that recruitment would continue to emphasize the service-providing rather than the knowledge-building functions of social work. This project further assumed that social work knowledge should be developed by social workers; and that, at least in the near future, persons with Master of Social Work degrees could and must take a major role in building and extending this knowledge through research. This latter consideration, in particular, was central to the broad goals of the project: identifying and maximizing research learning potential of students in social work Master's degree programs.

Specifically, this project sought to learn how, at the time of admission, schools of social work could begin to identify various levels of potential, flair for, and interest in doing research among students. At the same time, it tried to learn what were some of the teaching methods and course content most likely to maximize potential at various levels. In the process, it expected to test and refine an instrument that could be used both to identify levels of potential and to measure learning achieved.

In designing this training project, the writer and others involved were aware that many teachers do not view the research classroom course or courses and the experience in research provided students (often called the research project or practicum), as separate entities. For several

reasons, however, a decision was made to concentrate exclusively on the classroom courses in research at the five schools concerned.

First, success in any research experience such as a project or thesis was considered to be dependent on success in the course. Evidence for this is the almost universal requirement among schools of social work that students take at least one basic classroom course in research prior to any research experience. Second, the classroom course lent itself more readily to objective evaluations of the outcome of training than the project. Examinations using multiple choice questions were more likely to be used and acceptable to students as a basis for course grades, than in the project or thesis, where any grading was based on the narrative report itself. Third, findings from this pilot training plan for the research classroom course were expected to be useful in developing subsequent training plans for research projects and theses.

This pilot training project must be considered exploratory. It concentrated on study of methods of measuring potential for learning research among social work students and on study of the influences on student learning from existing differences in research teaching rather than from planned experimentation. Use of existing statements of content and objectives as data sources for differences in teaching resulted in some lack of detail on these topics. Consequently, findings are more general than specific. Nevertheless, it served its purpose by indicating that differences in potential could be identified, and by offering suggestions for changes in research teaching that could maximize this potential for each student.

Data on 353 students at the five participating schools were

obtained during 1965-66 prior to these students having any classroom training in research. Data on 263 students were obtained after they had completed their classroom work in research in 1966 and 1967.[3] Student attrition and some difficulty in obtaining some of the material at two of the schools were responsible for the loss, but the 90 students lost did not appear to effect the representativeness of the sample. Analyses of background data and test results showed that these 90 students were not significantly different at the 5 per cent level from those who completed the study.

The five schools, selected for their variety in size, geographical location, source of support, and length of accreditation, provided 11 teachers who taught a total of 14 sections.

The training plan exploited a number of naturally occurring differences among the five schools that influenced student learning. These included such factors as the relative emphasis on various aspects of content, differences in methods of teaching, and in learning experiences, variations in methods of evaluating learning achieved, the time and place of research in the curriculum, and teacher characteristics. As part of this training plan, two schools set up experimental sections where new content was introduced or new methods of teaching attempted, or both. Little or no change was made in the teaching at the other three schools.

Data on the objectives, methods of teaching research, course

---

[3]Data at the various schools were collected at different times because the research courses did not all begin and end in the same semester. At each school data were collected prior to any classroom work and subsequent to its completion.

content and other training details were obtained through interviews with a representative from each school and from the course outlines and statements of objectives, etc., prepared by each teacher.

Student characteristics prior to taking any research course work were obtained by questionnaires and a test. Two questionnaires obtained biographical data about students. A measure of students' attitude to and knowledge of research was obtained by means of a test that the writer had been developing during the last five years. This instrument, called the "Measurement of Attitudes and Research Knowledge," and referred to as the MARK, (by the first initial of each letter of the title) had been found in several studies at Tulane to predict which students did best in the research course, and which could benefit from special instruction.[4]

Further material on this instrument will be available in a companion monograph on this training plan, to be published approximately July, 1968.

Subsequent to students completing the beginning research course, but prior to any work on the thesis, research practicum or project, the instrument measuring attitudes to and knowledge of research (the MARK) was administered again. In addition, students were given an examination containing certain specialized and somewhat more difficult than average questions about research to determine what specialized learning had taken place. To get further information on teaching methods, at the same time, students completed a questionnaire evaluating the teaching

---

[4]Harris K. Goldstein, "Student Social Work Characteristics Related to Potential for Learning Research," 1965, mimeod material. Also see other papers on this instrument by various research apprentices of the writer produced during 1962-63-64, and available in mimeod form from the writer.

at each school and section and noted some of their general reactions to research. Final grades in the research course completed the evaluative data.

Analysis was guided by the objectives of the study. Using data obtained from the instruments administered after the course as criteria, determinations were made of the ability of the various pre-course instruments and questionnaires to identify potential for learning research. Changes in MARK scores, special test questions, and findings in the course evaluations, as measures of learning achieved, were then analyzed for their relation to the teaching variables.

In the material that follows, Chapter II presents some historical material about the profession's concerns with the teaching of research and some problems in and theory for identifying and predicting success in learning research. Chapter III explains how the pilot training plan was carried out and how the data used to evaluate its success or failure were obtained. Chapter IV provides some data on students prior to training. In Chapter V, information on the teaching variables that were part of the overall training plan is presented. Chapter VI discusses the adequacy of the instruments used to evaluate learning. Chapter VII gives the outcome of the training at the participating schools, together with an evaluation of results. Chapter VIII discusses further training and evaluation suggested by this project.

To preserve anonymity, all participating schools are designated only by number. The size of the sample at each school, the location, and other information that might identify schools has been omitted. Detailed tables on each school (identified by number only) on all variables analyzed in this study may be obtained from the writer by interested researchers.

CHAPTER II

STUDIES OF RESEARCH TEACHING AND SCIENTIFIC PRODUCTIVITY, AND THEIR RELATION TO STUDY METHODOLOGY

This chapter discusses some of the past and current concerns relating to the teaching of social work research. It presents findings of other studies aimed at identifying scientific productivity and creativity, and indicates why the particular methodology used in this study was selected.

The practical goals of this training plan differed in focus from some of the profession's previous concerns about the teaching of research. They did, however, have some similarity to attempts in other fields to learn what makes for success in research and to some current social work explorations of how to maximize learning for different levels of student ability.[1]

Previous Concerns With Research Teaching

At first glance, the concerns expressed most commonly in the literature about the teaching of social work research are not the ones on which this training plan was focused. Further assessment of these concerns, however, does show the relevance of this project to the profession's needs.

---

[1]At the Annual Program Meeting of the Council on Social Work Education in 1966 there were workshops on teaching the gifted and disadvantaged student.

The two topics most frequently mentioned in articles about the teaching of research are integration of the content taught in research with the remainder of the social work curriculum and methods for teaching the research project or practicum that will maximize student interest. At times, the former question has been expressed as a concern how to provide an educational plan in which an attitude of scientific inquiry and ability to use the scientific method in problem solving will permeate the entire curriculum. How to give students an appropriate research experience has been the focus of the second interest.[2]

The writer and others involved in this training plan recognized the importance of these previous topics. It seems clear, however, that integration of the research sequence and the remainder of the curriculum is related to how the research course is taught and how the rest of the curriculum is taught. The thesis of this training plan was that if better research teaching takes place, problems in integrating research with the rest of the curriculum and in developing and maintaining

---

[2]The most sustained concern relative to teaching research has been manifest in documents of the Council on Social Work Education. Review of CSWE workshop reports from 1954 on "Social Work Research," from 1957 on "Social Welfare Research and Demonstrations," and from 1964 on "New Approaches to Administration and Research in Social Work Education," all support the conclusion that integration of research and the teaching of the project have received major attention. Only one out of seven papers in "Selected Papers and Methods of Teaching Research in the Social Work Curriculum," published by the Council on Social Work Education in 1963 discusses the teaching of the research course. Four of these papers were about teaching the project, one about the post-master's program, and one about integrating the teaching of research and other practice methods.

interest in the practicum will be reduced.[3]

No material could be located in the literature related to the writer's concern for identifying and maximizing research learning potential. Correspondence with other faculty did locate two teachers with similar interests. Professor Irving Lukoff at the University of Pittsburgh, and Professor Merlin Taber at the University of Illinois both reported that they had been interested in the variations in knowledge that students brought to their research courses. Dr. Lukoff has done some beginning work on a test similar to the one developed by the writer to identify these differences. Dr. Taber modified Dr. Lukoff's test and did some further investigation with it. As far as the writer can determine, however, development of these two instruments is still in progress, and no reports describing even preliminary findings are available.

Dr. Lawrence Northwood attempted to enrich the training of certain students in the Master's program but his design differed from the training plan of this project in three ways: in the way students were selected for more intensive training, in the method used to provide this training, and in how the outcome was evaluated.[4] This study, carried on at the University of Seattle, selected students for

---

[3]Some current writings imply this though they do not state it explicitly. See R.M. Titmus, "The Relationship Between Schools of Social Work, Social Research and Social Policy," Journal of Education for Social Work, Spring, 1965, pp. 68-75, and E. Walsh, "Research and Teaching Casework," Journal of Education for Social Work, Fall, 1965, pp. 47-52.

[4]Lawrence K. Northwood, "Enriched Training in Research in the Master's Degree Program," Social Work Education Report, June, 1966.

additional training in research on the basis of variables which appeared to have face validity, including their stated interests, completion of one or more courses in behavioral science research, employment in research positions, and undergraduate grade point averages. Methods used to maximize the potential of these students included providing them with a classroom course in addition to the basic research course, and considerable additional work in a special practicum during the summer. Evaluation of students' performance was made on the basis of a paper, classroom participation, and a final examination (kind unspecified).

The relatively few reports of previous attempts to work toward objectives similar to this training plan, however, does not mean that there is agreement on how to identify differences in students' abilities nor even that there is any consensus on what the content, emphasis, or teaching methods in social work research should be.

The Research Workshops of the Council on Social Work Education previously cited, raise such questions as "How much statistics should be required, and how much statistics should be taught as part of the research course? Should the research course be taught in the first or second semester or in the first or second year? Should there be a full professor whose time is allocated to the teaching of research or should all faculty participate in this function?"

That definite answers are yet to be found is evidenced from a study of school catalogs. The wide variety of content mentioned, hours of credit given, the semester in which research is taught, and the experience and background of persons teaching the course among the more than 60 schools of social work is an indirect acknowledgement that the profession is far from having decided what is the best plan of teaching even a single

course, much less of integrating the research sequence with the rest of the curriculum.[5]

Theoretical and Empirical Bases for Predicting Success in Research

While there has been little study in social work of the factors within students making for success in learning research, or little investigation of how to teach research in a way that makes it most likely these factors can operate, there has been considerable investigation in other fields of factors influencing success in scientific effort. These efforts have generally been aimed at determining which factors relate to scientific creativity or which factors can predict research productivity subsequent to the completion of graduate work.

An excellent summary of theoretical formulations and empirical findings related to these topics is found in a report on the Utah Conferences on Creativity where workers from different disciplines in various parts of the country brought their knowledge together.[6]

While the objective of this training plan was not to identify or maximize potential for creativity, per se, among social work students, studies of the factors related to creativity were considered relevant to its goal. This relevance was based on the idea that factors influencing research creativity and productivity are similar to those that influence learning substantive content about research.

---

[5]Julanne Haspel, unpublished material, "Variations in Master's Research Teaching at Schools of Social Work," available in mimeographed form from the writer.

[6]Calvin W. Taylor and Frank Barron, Scientific Creativity, Its Recognition and Development. New York: John Wiley and Sons, 1963.

Theoretical formulations and empirical studies aimed at determining the factors influencing scientific creativity and productivity were found to deal with three kinds of variables related to intellect, personality, and culture, respectively.

Studies of productive scientists have been carried out by various methods ranging from clinical interviews and projective techniques, through empirically developed biographical inventories and factor-based tests. Various criteria of scientific creativity and productivity have been included, ranging from the numbers and kinds of research publications to supervisory ratings of personnel. From all of these various studies, a consistent picture of the personality and intellectual characteristics of a successful scientist emerges. The successful scientist possesses:

"1. A high degree of autonomy, self-sufficiency and self-direction.

2. A preference for mental manipulation involving things rather than people. Accompanying this is a distant or detached attitude in interpersonal relationships and preference for intellectually challenging situations rather than socially challenging ones.

3. High ego strength and emotional stability.

4. A liking for method, precision, and exactness.

5. A preference for such defense mechanisms as repression and isolation in dealing with affect and instinctual energies.

6. A high degree of personal dominance but a dislike of personally toned controversy.

7. A high degree of control of impulse, amounting almost to over-control; relatively little talkativeness, gregariousness, impulsiveness.

8. A liking for abstract thinking with considerable tolerance of cognitive ambiguity.

9. Marked independence of judgment, rejection of group pressures toward conformity and thinking.

10. Superior general intelligence.

11. An early, very broad interest in intellectual activities.

12. A drive toward comprehensiveness and elegance in explanation.

13. A special interest in the kind of 'wagering' which involves pitting oneself against uncertain circumstances in which one's own effort can be the deciding factor."[7]

This cluster of traits may require some modification to describe a successful social work scientist, possibly modification toward more interest in people and a change toward more interest in social relationships. Otherwise, it appears to have utility for identifying and predicting those persons who will likely have most potential for learning social work research.

A small number of sociological and cultural variables have been found to be associated with scientific productivity.

Students from rural areas and small towns, those whose families are in other than white collar occupations, and those with lower-middle class status tend to be most productive. Colleges and universities in the mid-west and far west, of moderate but not necessarily lowest cost, appear to produce most productive scientists. Student cultures characterized by humanism, breadth of interest, and reflectiveness were found to be associated with scholarly productivity, and cultures characterized by participation and aggression were negatively related to this productivity. Natural science productivity, on the

---

[7]Taylor and Barron, op.cit., p. 385.

other hand, was found associated with student cultures stressing scientism and aggression. Outstanding achievement in natural sciences does not appear in colleges where student cultures stress social conformity.[8]

Teachers who are most successful in encouraging students to learn research and to become productive scientists are those who are said to have such traits as demanding high standards of performance, assertiveness, and entrepreneurial vigor. Breadth of interest, administrative participation, and teaching zest, are additional assets.[9]

Identifying Research Potential in This Training Plan

In this training plan, the problem of identifying levels of potential for learning research might have been approached by examining the large number of intellectual factors already identified, and for which psychological tests have been developed. An alternate method would have been to use existing personality tests to learn the extent to which social work students possess each of the personality factors identified by various investigators.

Two problems were considered to preclude the use of either of these approaches. Either method would have required some criterion of success in learning social work research. No satisfactory criterion of known reliability was available; classroom grades were not considered suitable for this purpose. In addition, the writer has not found

---

[8]Donald L. Thistlewaite, "The College Environment as a Determinant of Research Potentiality," in Taylor and Barron, pp. 265-277.

[9]Robert H. Knapp, "Demographic, Cultural, and Personality Attributes of Scientists," in Taylor and Barron, p. 210.

standard psychological tests fruitful in making predictions about social work students.[10] The high degree of homogeneity among social work students, in terms of their ability and personality, reduces the predictive power of tests which predict satisfactorily in more heterogenous populations where the range of test scores is broader.

On the basis of the empirical studies referred to, the writer thus developed a simple questionnaire to assess biographical data in the familial, cultural, and educational area. This was supplemented by a measure of interests, on the grounds that personality variables as well as some background influences would be manifest indirectly through expressed interests.

The decision to place major emphasis on using the MARK in this study to identify potential for learning research was based on the following line of reasoning:

The writer assumed that in the course of social work students' life experiences prior to attending a graduate school of social work, the amount of knowledge they obtained about research and the attitudes they developed to it would be influenced by both their personality make-up and their intellectual abilities. Students with the intellectual ability and/or personality traits likely to facilitate success in learning research were considered likely to have learned more research over their life experiences and to possess more positive attitudes to it, than students who did not have these intellectual abilities or personality traits. While <u>opportunity</u> to learn research and to develop

---

[10]Harris K. Goldstein,"Prediction of Success in Schools of Social Work,"Summary of Research, George Warren Brown School of Social Work, Washington University, 1959-60.

positive attitudes to it could be considered an intervening variable in the equation above, the writer further believed that students with the intellectual ability and personality traits associated with success in learning research would, in the course of their life experiences, tend to find such opportunities. This latter assumption is supported by theories of vocational choice and by the experience of the writer. He has found that some students could not understand why they had scored so high on the MARK before taking the research course, since they had no background in research. Interviews with these students invariably brought out the fact that they had voluntarily sought out the kind of material which would tend to increase their knowledge of research and positive attitudes to it. For example, these students tended to read such literature as science journals, science feature articles in the newspaper and science fiction.

A test like the MARK that measured knowledge and attitudes was thus considered an indirect method of identifying those intellectual abilities and personality traits likely to influence potential for learning research. At the same time, the MARK provided an economical method of assessment of the outcome of this training plan, because it could serve as a criterion for what students learned in the research course. The MARK thus served a double purpose and provided: (1) a uniform method of measuring attitudes to research and what students had learned about research prior to taking the research course at all five schools; (2) whether any changes in their attitudes to and knowledge about research had taken place after they were taught social work research.

CHAPTER III

METHODOLOGY OF THE TRAINING PLAN

This chapter gives a summary of methods used to get data on students before and after they were taught research and to obtain information on teaching objectives, course content and methods used to implement student learning.

Method of Obtaining Data on Students Prior to Their Taking Research Course

Data obtained on students prior to their enrollment in the research course were obtained for two purposes that were in line with the overall objectives of the training plan. First, the aim was to obtain data considered likely, on the basis of material discussed in Chapter II, to predict different levels of potential for learning research. Second, some of the data was to be used as a base line from which to measure learning achieved.

Because the writer had been attempting to evaluate his own teaching since 1961, three previously tested instruments for gathering data were available for use in this pilot study. These instruments, details on which are given below, were as follows:

1. The first was a questionnaire which obtained identifying information on students' undergraduate college experience as well as material about their family and cultural background. In the remaining chapters of this report, this questionnaire is referred to as the "I.D.

Questionnaire" and the information obtained from it is called "I.D. data."

2. The second was a questionnaire which obtained certain information on students' interests, spare time activities and hobbies. This questionnaire is referred to as the "Interest Test" and information obtained from it is called "Interest data."
3. The MARK, a test measuring attitude to and knowledge of research was the third instrument. Data obtained from it includes an attitude score, a knowledge score, and a total score obtained by combining these other two scores.

The I.D. Questionnaire:

The I.D. Questionnaire was developed by the writer in 1963 to obtain data on students' educational and cultural background.

Items initially selected for this instrument were those that appeared to have face validity; others were suggested by the studies of success in scientific endeavor in other fields such as psychology, medicine, and physical science, referred to in Chapter II. Items of the face validity type were those that asked whether students had course work in statistics or research while an undergraduate, or those asking for a self-assessment on attitudes to and ability in mathematics and abstract reasoning. To these were added questions on demographic variables, such as size of hometown and geographical area of students' origin and cultural variables such as social class.

On the basis of item analyses carried out in 1964 and 1965, certain of these items which did not correlate with final research course grades

at Tulane were eliminated. Other items were modified and some new items added.[1]

In the form used in this pilot training plan, this I.D. Questionnaire consisted of 32 pre-coded items. It was administered to students as a group at each of the five schools before they had any course work in research at the school of social work. In general, students completed this questionnaire in about 20 minutes.

The Interest Test:

This test was originated by the writer in 1963, at the same time as the I.D. Questionnaire. Because so little was known regarding what interests might relate to different potentials for learning research, it purposefully tried to cover a broad range of interests. Use of a broad rather than a specific test was supported by what is known of other interest tests, where in some cases, rather unexpected correlations were found between certain interests and success in a field. The test format asked the student to check those activities he had "spent some time on, either as a paid worker or as a voluntary or recreational activity in the last five years." A list of activities was provided.

Originally, this test contained 60 items. In 1964 and 1965 other items were added similar in content to those original ones that correlated with research course grades. At the same time items

---

[1]Mimeographed papers by Douglas Radabaugh, Julanne Haspel, Faye Weckel, and Eugene Hadley, Student Research Apprentices, who worked with the writer on these analyses are available to interested persons. Only those items that correlated at the 5 per cent level with research course grades or those supported by other theoretical or empirical studies were retained on the questionnaire.

showing less predictive value were eliminated.[2]

This test was administered immediately following the I.D. Questionnaire to students at each of the five schools as a group. In the form used, it consisted of 64 interest items to be answered "yes" or "no" and required about ten minutes for students to complete.

The "Before" MARK:

The writer began work on this test in 1961 at Washington University. Work on this test was stimulated by the belief that students brought to the research course a far wider range of knowledge and attitudes than they brought to other courses, a range that posed special problems in teaching research. By administering the MARK before the research course, the writer attempted to identify various levels of student potential for learning research as measured by course grades on objective tests. By administering it again after the course, test scores served as an indicator of the amount of learning achieved.

In 1961, the original MARK consisted of approximately 30 items taken from multiple choice final examination questions of the writer. Students' pre-course scores on this original test confirmed the writer's belief about the relatively wide range in knowledge that students brought to the course.[3] Some students scored sufficiently high so that had this been a bona fide final examination, they would have been given a grade of "B+." Others scored so low that their score was less than

---

[2]See papers previously referred to by student apprentices.

[3]A study by the writer of variations in casework knowledge prior to students taking any casework courses showed considerably less variation in the casework area than in research. See "The Tulane Assessment Scale for Caseworkers," Tulane Studies, Vol. VII, by the writer.

15 per cent of the score made by the "B+" students. (On examinations at the end of the course, even the failing students generally score at least 50 per cent as much as "B+" students.)

In succeeding years at Tulane, this original test was modified and enlarged on the basis of item analysis, or correlations of each item with every other item, and with grades on objective tests at various points in the research course. In 1964, certain items considered to measure attitudes as well as knowledge were added. These original "attitude" items were selected on the basis of the writer's inferences which attitudes appeared to make a difference in student learning. In 1964 and 1965, the predictive value of these items was also tested by item analysis.

In the form used in this pilot training project, the MARK had a total of 59 pre-coded, multiple choice questions, of which 36 were considered to measure "attitudes" and 23 to measure "knowledge." Attitude and knowledge items were distributed at random throughout the first 40 items of the test. The last 19 items were "knowledge" items. The MARK was administered to students at each school in a group before they had taken any research courses.

In the material that follows, this form of the MARK is referred to as the "before" MARK to differentiate it from that slightly different form administered after students finished their research course.

Method of Obtaining Data on Students Subsequent to Their Completing the Research Course

At the end of this pilot training plan, four kinds of information were obtained. Three of these also capitalized on previously developed instrumentation of the writer.

1. A slight modification of the MARK called the "after" MARK provided knowledge about and attitudes to research.

2. A group of questions found by experience to be unusually difficult for Master's level students were used to obtain information on any learning above and beyond that usually expected at the MSW level. These questions, objective in form, are referred to as "Special Test Questions," or the STQ.

3. A questionnaire asked for students' reactions to the teaching and content of the course. This is referred to as the "Course Evaluation."

4. Grades in research courses, based on whatever method each teacher used in assigning grades, were obtained.

The "After" MARK:

With slight modifications, the MARK was used to obtain an attitude score, a knowledge score, and a total score that would show changes in these three variables after the students completed their research course work. The "after" MARK differed from the "before" MARK in only four items. In the "before" MARK, these items had dealt with student expectations about the research course; in the "after" MARK these items were changed so that they asked about student experiences rather than expectations. For example, before items asked about apprehension about research, or about the expected application of the students' learning to their practice. Corresponding items in the after form asked if the course had caused as much anxiety as expected, or if application of learning was more or less than expected.

The "after" MARK was given the students at Schools 1,2,4, and 5 in a group administration immediately after students finished their research course. At school 3, students were asked to take the MARK home

with them for the summer and return it by mail. Some attrition occurred at each school, so that after data was not obtained on all students on whom before data was available. For details, see Chapter VII.

The Special Test Questionnaire:

In 1965 the writer developed some questions aimed at learning whether students who had been placed in a "special" section and exposed to certain additional concepts (see Chapter V for details) had learned any additional content. Analysis of student responses to these questions at Tulane showed them to be of more than average difficulty. No student received a score of more than 80 per cent of the items correct, and the average score was about 50 per cent correct. These "special" questions were used in this pilot training project to evaluate if any learning, above and beyond that generally expected, had taken place.

In the form used to evaluate the success of this project, there were 20 of these questions, all multiple choice. They cut across research methodology by presenting 20 possible procedures to be used in a predictive research study. Four possible outcomes of each procedure were provided. These were (1) whether following the procedure would produce an increase in the reliability with which the predictor variable could be observed, but have no influence on the validity of predicting the expected relationship; (2) whether the procedure would increase the accuracy of prediction of the criterion variable, but have no influence on the reliability with which the predictor variable was observed; (3) whether the procedure neither increased the reliability or validity (including those whose influence was unknown or which were likely to decrease reliability and validity) of the prediction; and (4) whether the procedure would both increase the

reliability with which observations of the predictor variable could be made and at the same time increase the validity with which the prediction could be determined. Students were asked to select which outcome of each procedure was most likely.

At school 5, these questions were administered as a part of the final examination. At schools 1,2, and 3, they were given in the Fall of 1966. They were not administered at school 4.

Course Evaluation:

Since 1957, the writer had been using a questionnaire administered to students at the last class session of the research course, to obtain students' reaction to teaching methods in general, to specific aspects of content, and to get students' general reaction about research and the learning they had achieved.

Like the other instruments in this study, this questionnaire was modified from year to year. Modifications were made, however, on the basis of logical rather than statistical analysis. The instrument was updated in keeping with changes in the content and teaching methods of the course taught by the writer. It is believed these changes made it more inclusive of the points about which a teacher would be concerned, but this inclusiveness was not evaluated specifically for the five schools in this project.

In the form used as part of the evaluation of this pilot training project, this questionnaire consisted of 30 items. It included inquiry as to how helpful students had found the research course to be, how much interest they had in research after taking the course, their reaction to teaching techniques, methods and content, and their own estimates of their knowledge gained. A number of questions inquired specifically

about which parts of the content was most difficult or easiest, best and least liked, least and most useful in practice, taught best or worst, and requiring most and least study time.

To increase the probable validity of responses, students were asked to complete this form anonymously. To permit some comparison of student course evaluations with grades, students who had received their final course grade were asked to put it on the course evaluation form.

Teachers at the five schools, on their own initiative, had students complete this form in different ways. At school 5, students completed the course evaluation the last day of the course when they had already received their course grades; at school 2, student grades were not ready and students completed the course evaluation the last day of the course without knowing their grades. At school 4, the evaluation was completed at the same time as the "after" MARK. At schools 1 and 3, the evaluation was completed in the Fall of 1966 when students had returned from the summer.

Course Grades:

Although grades earned in courses may be of dubious reliability and validity as a measure of knowledge acquired, they are a readily available and customary method of assessing students' learning. Therefore, course grades were used in this project as one method to determine learning.

A theoretical 10 point scale was used to convert letter grades to numerical ones with A+ = 9, A = 8, A- = 7, etc., down to C- = 1 and D, E, or F = 0. No school made use of the full ten point scale, thus introducing an extraneous variable into the analysis of results. Two schools' grades ranged over eight scale points, one school used five

scale points, while two schools included only three and two scale points, respectively, grading only A, B+ and B, or A and B.

Methods of Obtaining Data on Teaching Variables

Data on teaching variables were obtained by content analysis of the sources listed below.

School Catalogs:

School catalogs provided information on three variables, including the hours of credit for research, the period in the school year during which research was taught, and some information about the content of the course.

Written Material from Teachers:

Written material prepared by the 11 teachers whose classes participated in this study covered two variables. These were course objectives and the relative emphasis on various aspects of course content.

Correspondence and Interviews With Assistant Project Directors:

Material obtained by correspondence and interviews with the Assistant Project Directors at each school included data on ten variables. These were: any experimental objectives, content or methods; the evaluation and examination method used at the school; the learning experiences provided students; the background of the teacher including the level of the teacher's training and the teacher's field of specialization; the method of sectioning the course; and whether special classes were available for above average students or remedial classes available for those with difficulty.

CHAPTER IV

STUDENT CHARACTERISTICS PRIOR TO TRAINING

This chapter presents information on student characteristics prior to taking the research course. A comparison of this sample with students in the Pins study of students in social work schools in 1960 is made first. A profile of the typical student is then presented, followed by sections in the chapter on each of the variables on which data were obtained prior to training. In the interest of brevity, the I.D. material and data on interests presented in this chapter deal with all schools as a whole. Identifying data and interest variables found to predict differences in potential for learning research are compared school by school in Chapter VII.

Because the MARK is the major predictor in this project, "before" MARK scores are provided in more detail than the I.D. and interest data, with school to school comparisons being made.

Comparison of This Sample With Students in Other Schools of Social Work

Six identifying variables on which information was obtained for this project were categorized identically with that obtained in the Pins Study of 1960.[1]

---

[1]Arnulf M. Pins, "Who Chooses Social Work, When and Why," Council of Social Work Education, 1963.

While information on these six variables was obtained because of its possible value for identifying potential for learning research, the similarity of these variables to the Pins Study provided an opportunity to determine how representative this project sample was of social work students throughout the country. Such a comparison, of course, assumes there has been no change in student characteristics from 1960, the date of the Pins Study, to 1965, the time at which these data were collected.

The six variables common to the training project and the Pins Study were parents' income, father's occupation, father's education, mother's education, size of home town and the type of college or university from which student received an undergraduate degree.

This project sample scored somewhat above the Pins sample in respect to father's occupation, and the education of both father and mother. Significant differences at the 5 per cent level were found. Tests by chi square showed no significant difference between the distributions in the Pins Study and this project on parent's income, size of home town, and type of undergraduate college.

Comparison of this group and the group in the Pins Study by X^2 on grade point average, showed a significant difference at the 5 per cent level. The Pins group had a slightly higher undergraduate grade point average.[2]

Typical Student Profile

The father of the typical student was employed in a professional or managerial occupation, earned less than $7,000 per year, and neither

---

[2]Because grades were not obtained in identical categories in the two studies, this comparison necessitated some combining of categories in both studies.

he nor the mother was likely to have attended college.

The size of the students' "home town" was most apt to be a city above 250,000 in population. Cultural resources of "home towns" were generally said to be "far below average." In contrast to this characterization, most students reported their own cultural background was somewhat broader than average.

Education at a public university having a student body larger than 2,500, located in the southern United States was most characteristic of the group involved in this pilot training plan. The typical student was about as likely to have had a course in research, statistics, logic or scientific method as not to have had such a course. If he had taken any of these subjects, this was more likely because they were required than as an elective.

As a rule, the typical student had avoided contact with mathematics and abstract reasoning in the past. Of those who had some experience with mathematics, about an equal number reported it had been pleasant as said it was unpleasant.

About one-half of the students had been exposed to some classroom instruction, not always an entire course, related to research, usually in statistics.

About as many student interests involved contact with people as did not. Many of the more frequently reported interests were oriented to cultural goals and could be considered to be more characteristic of women than men, such as listening to music, going to plays, reading poetry, etc.

On the "before" MARK test, the typical student tended to give correct responses to about one-half of the knowledge items and to offer

responses considered desirable to about the same proportion of attitude items. The modal total score was 29; the modal attitude and knowledge scores were 12 and 17, respectively. That the knowledge and attitude items appeared to measure different aspects of the student is evidenced by the rather low correlation between them, .14.

I.D. Variables

Parents' Income, Occupation, and Education:

Almost one-half of the students came from homes where family income was less than $7,000. (The average income in the United States in 1965 was about $7,300.) Exactly one-half of the fathers were either professionals, proprietors, or in managerial occupations. Over two-thirds of the students' parents, both fathers and mothers, had no training beyond high school.

Cultural Background:

Students tended to come from urban backgrounds. About one-half were from cities of 250,000 or above.

Although slightly more than one-half of the students considered the adequacy of cultural resources in their home town below average, almost three-fourths of the students considered their personal cultural background broader than average. Where they received this additional cultural background is not clear.

In answer to a question about the amount of professional literature in their homes, such as medical magazines, or business and engineering journals, three-quarters of the students reported at least an occasional periodical or book of this kind. Only about one-half of the students, however, reported reading this material as much as occasionally.

Educational Background:

Slightly less than half of the students, four out of ten, had gone to a public university and about one-half had attended a large college of over 2,500 students. The largest proportion of students, or about four out of ten, had gone to a southern college or university (perhaps reflecting the fact that there were two southern schools in this project). The next largest proportion had gone to an undergraduate college or university in the northern United States. These three questions, as might be expected, produced most variation from school to school.

About one-half of the students had received credit for an undergraduate course with content covering in whole or in part statistics, research, logic, scientific method, or psychological tests. Two or more hours of undergraduate credit in statistics had been received by more than half of the students. About two-thirds of those who took courses in research, statistics, or scientific method, had done so because it was required. About half of those who took courses in logic and tests and measurements did so as an elective. Three-quarters of the students had this kind of training within five years of attending a school of social work and about half of them had taken these courses within two years prior to entering graduate school.

Background in Mathematics and Abstract Reasoning:

About half of the student group reported they usually avoided contact with math but approximately a third said they were neutral to, and about one-fifth reported a liking for this field. Students' attitude to math does not seem connected with their past experience, at least as reported in this study. Only one in five reported unpleasant

past experiences with math and one in six reported pleasant experiences, although about half had both pleasant and unpleasant experiences. The proportion who said they considered themselves to have about average ability in math (about half of the students) was about the same as those who liked it or were neutral to it.

About as many students characterized themselves as preferring abstract reasoning as avoiding it. About six out of ten, however, said they preferred doing to thinking, and only 5 per cent preferred thinking to doing. The others were about equally divided on this question.

Spare Time Activities and Interests

The most common interests of the students in this project did not appear different from interests of other young people their age.

The ten spare time activities and interests reported most often by 65 per cent or more of the students were as follows:

Listening to music	93%	Sightseeing	75%
Reading news items	87%	Recreation with children	72%
Reading non-fiction	83%	Going to plays	71%
Bull sessions	80%	Swimming	71%
Travel	77%	Attending church regularly	65%

Reported by 45 per cent or more of the student group were these:

Walking	64%	Reading poetry	52%
Cooking	63%	Hiking	52%
Visiting art museums	57%	Tennis	48%
Reading history	54%	Planning community work	47%
Singing	53%	Church social groups	46%

Activities and interests checked by from 10 to 20 per cent of the group were as follows:

Sailing	19%	Political campaigning for a candidate	17%
Painting	19%	Auto mechanics	14%
Hospital volunteer work	18%	Ceramics	13%
Studying the stock market	18%	Woodworking	12%
Folk dancing	17%	Betting on horse races	11%

Activities and interests least typical of the student group and checked by less than 10 per cent of the sample were:

Boy or girl scout leadership	10%	Chemistry	9%
Shell collecting	10%	Physics	5%
Flower arrangement	10%	Philately	5%
		Gem collecting	2%

Background in Research

With regard to students' background in research as measured by the "before" MARK, considerably more similarity than difference was found between each of the five schools. This similarity appears sufficiently close to say that students at the schools were matched so far as attitudes to research and knowledge of research are concerned.

This similarity in background in research is shown by the frequency distribution on "before" MARK scores in Tables 1 and 2 and the lack of significant difference between schools on both mean and median scores.

While there is a significant difference at the 5 per cent level between the mean scores of the lowest schools, 1 and 2, and the highest school, 4, this is barely significant and it is not significant at the

one per cent level. Differences between other mean scores were not significant for any of the other schools with regard to total scores, attitude scores or knowledge scores.

TABLE 1

MEAN SCORES ON "BEFORE" MARK TEST BY SCHOOLS

School	Mean Total	Stand.Dev.	Mean Attitude	Stand.Dev.	Mean Knowledge	Stand.Dev.
1	28.1	5.4	12.4	2.7	15.7	4.5
2	28.1	4.5	12.4	2.2	15.8	4.1
3	30.5	6.4	12.8	2.2	17.7	5.5
4	30.8	3.5	12.6	2.4	18.3	2.9
5	28.2	5.1	12.8	2.4	15.4	4.0

TABLE 2

MEDIAN SCORES ON "BEFORE" MARK TEST BY SCHOOLS

School	Median Total	Median Attitude	Median Knowledge
1	28.0	12.0	16.0
2	28.0	13.0	15.0
3	29.0	13.0	18.5
4	30.0	13.0	19.0
5	29.0	12.0	16.0

CHAPTER V

THE TEACHING VARIABLES IN THE TRAINING PLAN

Differences and Similarities in Objectives, Methods and Content

This chapter describes objectives sought at each school in the beginning research class, during this project, and the content and methods used to reach these objectives. It includes a statement of any experimentation by the schools. Because the selection of schools attempted to achieve maximum variation in school size, geographic location covered, source of support and length of accreditation, similarities among the five schools should be considered a finding rather than an outcome of the selection process.

Whenever possible, objectives and the relative emphasis on each within a school were obtained directly from written statements currently in use at respective schools. Early in the training plan, in September 1966, Assistant Directors from all schools, except school 2, met at Tulane and discussed the relative emphasis at their respective schools on the various objectives listed below. These discussions permitted a consensus on which schools had least and most emphasis on each objective.

Relative emphasis on course content was determined from statements in school catalogs, and by comparing these with class outlines and schedules when these were available. Inferences from this data were checked with Assistant Directors at the meeting given above, until

consensus was obtained on the relative emphasis given to various aspects of content at each school. School 2 later provided in writing information on objectives and content.

Relative emphasis on teaching methods, methods of evaluating learning experiences, and teacher characteristics were obtained from Assistant Directors. The lack of specific detail on objectives and content precluded analysis of possible relationships between emphasis on teaching specific content and change scores for certain MARK items that could be expected to reflect this teaching. Nevertheless, the use of global and general divisions of objectives and content was considered to maximize the reliability of decisions about relative emphasis on these from school to school.

A detailed written statement of objectives and content for each of the three sections was not submitted by school 4. Conclusions about school 4 are derived from general information obtained through the Assistant Director for that school.

All of the information obtained from the five schools shows considerable agreement on the objectives of the research course and almost as much agreement about its content. There is considerable difference, however, in the emphasis given various aspects of content, and in the teaching methods used and learning experiences provided.[1] Schools are

---

[1]One of the difficulties in collating and comparing material concerning course objectives is that content of written statements on these topics makes use of a high degree of abstract terminology. Another problem is the tendency to state objectives in terms of learning experiences. For example, the writer was not sure whether the objective "understanding the application of the problem-solving methods to professional problems" is the same as "understanding of research as a problem-solving method" or "understanding the scientific method of problem-solving." Interviews with representatives of schools followed by some correspondence with individual teachers were used to reduce errors as much as possible in this step of the analyses.

most different from each other in methods of giving assignments and of evaluating learning, in teacher characteristics, and in the time and place of the research course in the curriculum.

As far as could be determined from available information, there appears to be enough similarity in objectives and content so that the difference in other variables found may be looked at in terms of their influence on learning achieved in the course.

Objectives:

Courses at all of the schools share the goal of aiding the student to help clients better; the development of knowledge for understanding and interpreting research studies; and all seek to some extent at least, to prepare the student to carry out an actual research project or practicum.

A clear difference in emphasis on these objectives emerges from content analysis of schools' statements and from discussions with representatives of schools.

School 5 places somewhat more stress than any other school on the course enabling the student to help clients. There is a further difference on the emphasis in understanding and interpreting research studies. Schools 2 and 3, and possibly 4, place more emphasis on this than schools 1 and 5. Schools 1,2,3, and 4, place considerably more emphasis on preparing a student to carry out a research project, than school 5. These differences are summarized in Table 3 that follows.

TABLE 3

COURSE OBJECTIVES

	Schools with Most Emphasis	Schools with Least Emphasis
Preparing Students for Practice in Service Methods	5	1,2,3,4
Developing Knowledge and Skills in Evaluating Research Studies	2,3,4	1,5
Preparing Students to Carry Out a Research Project	1,2,3,4	5

Methods:

Schools with similar emphases on objectives were found to use similar teaching methods, to offer similar learning experiences, and to evaluate learning alike; schools with different emphasis on objectives were different in regard to these other variables. Schools 1 and 5 were most likely to be together, or at least most likely were different from schools 2,3, and 4. Schools 1 and 5 made most use of lecture and demonstration by the teacher and gave more tests of an objective nature. Others used seminar or class discussion methods and written narrative type assignments more often.

Tables 4, 5, and 6 which summarize the comparisons among the five schools on these topics should not be interpreted as meaning that a school fails to engage in a particular teaching method, that it does not provide a particular kind of learning experience, or that it fails to use a particular type of evaluation. The tables also should not be

interpreted to indicate any relative emphasis within a particular school on one or more of the points below. In addition, the conclusions below are presented in this general form rather than in any more specificity as a means of maximizing their reliability.

TABLE 4

METHODS OF TEACHING

	Schools with Most Emphasis	Schools with Least Emphasis
Lecture	1,5	2,3,4
Questions from class answered by class	2,3, and possibly 4	1,5
Questions from class answered by teacher	1,5	2,3, and possibly 4
Demonstrations by teacher	1,5	2,3
Exercises in class	3,5	1,2,4
Exercises at home	1,2,3,4	5

TABLE 5

METHODS OF EVALUATING LEARNING

	Schools with Most Emphasis	Schools with Least Emphasis
Objective tests	5,2	1,3,4
Papers (less scope than term papers)	1,3,4	2,5

TABLE 6

LEARNING EXPERIENCES

	Schools with Most Emphasis	Schools with Least Emphasis
Listening to lectures	1,2,5	3,4
Class discussion	1,3,4	2,5
Reading research studies	1,2,3	4,5
Writing or oral analysis of studies or reports	2,3	1,4,5

Content:

The emphasis on content, shown below, somewhat follows the emphasis on objectives and methods.

TABLE 7

CONTENT OF RESEARCH COURSES

	Schools with Most Emphasis	Schools with Least Emphasis
Statistics or Quantitative Methods	1,2,3	4,5
Substantive Knowledge of Past and Current Social Work Research	2,3	1,4,5
Phases of the Scientific Method[2]	1,4,5	2,3
Techniques and Instrumentation[3]	2,5	1,3,4

[2]Phases of the scientific method included such content on formulating hypotheses, selection of data sources and observational units, methods of data collection and research design.

[3]Techniques and instrumentation included such topics as how to make rating scales, how to test for reliability and validity, etc.

Teacher Characteristics:

Schools 1,4, and 5 had the most experienced teachers. These teachers were in general also the teachers with the most advanced training, and those whose training was in social work or in a field closest to social work.

TABLE 8

TEACHER CHARACTERISTICS

	Schools Where Teachers Had Most	Schools Where Teachers Had Least
Education and Training	1,4,5	2,3
Teaching Experience	1,4,5	2,3
Background like Social Work	1,2,5	3,4
Full Time Work	1,4,5	2,3

Miscellaneous Variables:

Miscellaneous variables such as hours of credit, and how early in the students' educational experience the research course occurred, etc., were less patterned than the other teaching variables mentioned above.

TABLE 9

MISCELLANEOUS VARIABLES

	Schools With Most	Schools With Least
Hours of Credit	2,5	1,3,4
Earliness of Occurrence of the Course in School Curriculum	1,2	3,4,5
Size of Class	1,5	2,3,4
Special Remedial Work	1,2,5	3,4
Later Elective Hours of Work Available	2,3,5	1,4

Experimentation Introduced as Part of the Pilot Training Plan

As part of the design of this pilot training plan, two schools, 2 and 5, introduced some experimentation.

At school 2, students were divided into four approximately equal sections. High scoring students on the "before" MARK were placed in Section A; low scoring students were put in Section B. Sections C and D each had a mixture of high and low scoring students. Statistics was added to the course content for students in Section A; students in Section B had no instruction in statistics; and Sections C and D received the usual content and emphasis, which included some statistics.

At school 5, the top 20 per cent of the students, according to "before" MARK scores, were put in Section A. Other students were assigned at random to Sections B and C so that each had about 40 per cent of the students.

In addition to the difference in size and students' test scores, Section A was different than other sections in two other ways: (1) There was considerably more discussion and interaction among the students themselves and less lecturing by the teacher. (2) Additional content was introduced that dealt with conceptualization, and specialized research techniques and measurements, content often considered more suited for advanced research courses.

At schools 1,3, and 4 student assignment to various sections was at random.

CHAPTER VI

EVALUATION OF INSTRUMENTATION AND DATA COLLECTION

Most of the discussion on the evaluation of instruments in this chapter deals with the "before" and "after" MARK, since this instrument played dual roles of major importance in the study. With course grades as the criterion, the "before" MARK was used both to identify various levels of student potential and to predict various levels of learning achieved. The "after" MARK served as a measure common to all schools to validate course grades, and also, when compared with the "before" MARK, to show differences in learning that had been achieved.

Evaluation of the MARK includes studies of its reliability, validity, homogeneity, difficulty, and ability to discriminate. The construction of the other instruments, with the exception of the special test questions, did not permit quantitative evaluations of reliability. Comments on the reliability of these instruments is thus based on logical analysis; tests of validity and discrimination were handled conventionally.

The MARK

During the six years the MARK was under development, its content was changed several times to increase its power to predict course grades at Tulane. Changes were not necessarily aimed at making it more representative of content taught either at Tulane or at other schools. Analysis of the content of the MARK used in this project, however,

according to the categories used in Chapter V to describe the content of the beginning research course as taught at the various schools of social work, indicates a fairly balanced sampling of the four content areas described.

Ten items on the MARK required knowledge of statistics or quantitative methods; six items related to knowledge of findings of past and current research; twenty-one items asked for knowledge of various phases of the scientific method; and seven items referred to techniques and instrumentation. The remaining fifteen items were all attitude items that did not appear to fall into any of these categories.

This relatively greater emphasis on the general aspects of the scientific method and the relatively lesser emphasis on substantive knowledge of findings from research studies, and techniques and methods, represents the relative content emphasis in the beginning research course at Tulane. Though this distribution of the content measured by the MARK did not necessarily fit so closely the content emphasis at the other four schools, it did not significantly effect results. When the Tulane students were removed from the sample, for example, and analysis made of the other four schools, findings at the other four schools remained consistent with those from Tulane.

The ten items about statistics or quantitative methods included seven questions requiring knowledge of terminology such as "parameter," "mean," "correlation," etc., and three questions on statistical procedures such as how to make a frequency distribution or the function of statistics in research.

The questions on knowledge of past and current research findings asked for general information about the findings of social work research

rather than specific knowledge obtained by any particular study. These questions covered such topics as whether social work research appeared to provide definitive or limited answers to problems in the field, how helpful social work research findings appeared to be in guiding practice, or how one evaluated the findings of a research study, etc.

The twenty-one questions on the phases of the scientific method included general inquiries about the method as a whole and specific questions about certain aspects. Examples of general questions were items that asked why the scientific method was considered self corrective, and whether research in social work was different from that in other fields. Specific questions about phases of the scientific method included inquiries into what was the first decision to be made in a research study, or what criteria topics for a research study must meet.

Questions on techniques and instrumentations were specific "how to do it" items. Examples of these were items asking how one differentiated abstract and concrete concepts, what guides were used in deciding how many classifications should be made from a given number of observations, how the validity of observations was determined, etc.

In other work on the MARK, the writer has found that test items which appeared conceptually to cover similar content did not predict equally well. Because this finding leads to some question whether items which appear conceptually similar are actually measuring the same thing, no MARK subscores according to the above breakdown of items by content were developed. In further work on the MARK, factor analysis will be carried out. Subscores will then be computed for items that are found to measure common factors and the relation of these subscores to variation in content at the five schools will be studied.

Reliability--"Before" and "After" MARK:

The reliability of the "before" MARK of .62 is considered to show a satisfactory consistency in measuring potential for learning research. The reliability of the "after" MARK as a measure of what was learned in the research course is even more consistent, .75. The reliability of the "before" MARK was somewhat less than the reliability of its parts; the reliability of the "after" MARK on the other hand, was somewhat higher than the reliability of the after attitude scores, but somewhat lower than the reliability of the after knowledge scores. These findings are shown in Table 10.

TABLE 10*

RELIABILITY OF "BEFORE" AND "AFTER" MARK SCORES

Test	Reliability	Test	Reliability
TOTAL BEFORE	.62	TOTAL AFTER	.75
Attitude Before	.66	Attitude After	.57
Knowledge Before	.76	Knowledge After	.84

N = 263

*All reliabilities are based on Pearsonian correlations carried out using the split-half method of correlating scores on even numbered items with scores on odd numbered items. All are corrected for the number of items, using the Spearman-Brown formula, so that reliabilities of attitude and knowledge items are those that would be found if there were 59 attitude and 59 knowledge items.

The increased reliability of the total scores and knowledge scores after students had taken the research course is what would be expected on the grounds that there would be less guessing on the various items after students had studied research. The reason why the reliability of attitude scores decreased after taking the research course is unknown. One

course in research may upset some previously formulated student attitudes about research but not necessarily help students to formulate others to take their place. Thus, at the end of one research course, student attitudes may be somewhat in a state of flux. If this is true, then studies of attitudes after the practicum that follows the research classroom course might show more consistency.

As might be expected, school to school variation in the reliability of the MARK occurred to some extent. These could be largely accounted for by the size of the sample at the different schools. Most variation from school to school occurred in the reliability of before attitude scores and least variation in after knowledge scores.

Item Difficulty and Heterogeneity "Before" MARK:

Some indirect evidence of the potential of the "before" MARK to identify various levels of potential for learning research can also be found in the relative difficulty of items and in the heterogeneity of the content.[1] The distribution of items by difficulty, as shown in Table 11 below, indicates the high variation in the difficulty of both attitude and knowledge items.

---

[1]Tests with items of varied difficulty and heterogeneous content have been found to be most successful in identifying various levels of potential for learning the content in other subject areas. See Lee J. Cronbach, Essentials of Psychological Testing, New York: Harper and Brothers, 1960.

TABLE 11

LEVEL OF ITEM DIFFICULTY--"BEFORE" MARK

Level of Difficulty as Shown by Per cent of Students Who Missed the Item	Number of Items at Difficulty Level Shown		
	Total Test	Knowledge Items	Attitude Items
80 - 99%	10	7	3
60 - 79%	12	9	3
40 - 59%	15	8	7
20 - 39%	13	8	5
0 - 19%	9	4	5
Total	59	36	23

The heterogeneity of a test may be measured by the number of inter-correlated items. Items having little or no inter-correlations may be assumed to measure different things.

The heterogeneity of the "before" MARK is shown by the fact that less than 6 per cent of the inter-item correlations are significantly greater than zero at the 5 per cent level. At the 1 per cent level, less than 1 per cent of the total inter-item correlations are significant.

The size of inter-item correlation also indicates relatively heterogeneous content. The highest inter-item correlation was only .42, and out of a total of 710 possible inter-item correlations, only four were above .30.

Discrimination:

The possibility of the "before" MARK identifying those students with different potentials for learning research is a function of its capacity to discriminate among students.

Over 80 per cent of 48 out of the 59 items on the "before" MARK discriminated among the students in some way.[2] As one might expect with a test given the students before they took the course, most items discriminated the top third from the other groups of students and fewest discriminated the bottom third from the other two-thirds, as shown by Table 12 below.

TABLE 12

ITEM DISCRIMINATION--"BEFORE" MARK

	Discriminated Top Third From Middle Third and Middle Third from Lower Third	Discriminated Top Third From Other Two-Thirds	Discriminated Lower Third From Other Two-Thirds
Number of Items	16	24	8
Number of Attitude Items	4	12	3
Number of Knowledge Items	12	12	5

Validity:

Support for the validity of using the MARK as a measure of learning achieved in the research course is found in the relatively high correlation (.56) of the "after" MARK with the usual measure of research learning, course grades, and from the discrimination of its items, their heterogeneity and level of difficulty.

[2] To determine the discrimination of MARK items, the students were divided into three groups according to total scores, the top third, the middle third, and the lower third. The discrimination of items shown in Table 12 indicates those items that discriminated the three groups from each other and those items that discriminated one group from the other two groups.

Since obtained validities are influenced by the reliability of both of the variables that are correlated, in Table 13, below, both the obtained validity, and the validity corrected for the lack of reliability of the inter-correlated variables are shown.

TABLE 13

VALIDITY OF "AFTER" MARK SCORES AS SHOWN BY CORRELATION WITH RESEARCH COURSE GRADES

	Obtained Validity	Corrected for Attenuation
TOTAL AFTER	.42	.56
Attitude After	.27	.47
Knowledge After	.41	.49
N = 263		

In making the correction for attenuation (column two above), it was assumed that the course grades had the same reliability as the MARK. Because many course grades are based on term papers or narrative examinations, it seems likely that most course grades have a lower reliability than the MARK. The correction made is thus on the conservative side and may tend to underestimate the true validity of the MARK. At the school that used wholly objective tests to evaluate learning achieved, for example, the correlation of "after" MARK total scores with course grades after correction for attenuation was .77.

School to school variations in the validity of the MARK as a whole were found also, largely related to the size of the sample at different schools and whether the grading was based on objective tests, short answer examinations, or on papers.

The relatively lower validity of the attitude items may be due to the fact that they are actually of lower validity. This may also be due to the fact that the attitudes which the writer considered were likely to be associated with success in research were not necessarily those which were so associated. Further study of attitude items under another scoring system will be discussed in the forthcoming material on the MARK.

Item Difficulty and Heterogeneity--"After" MARK:

There was almost as high a variation in difficulty of the "after" MARK items as for the before form. Changes as might be expected are toward a lower level of difficulty as shown in Table 14.

TABLE 14

LEVEL OF ITEM DIFFICULTY--"AFTER"MARK

Level of Difficulty as Shown by Per cent of Students Who Missed the Item	Number of Items at Difficulty Level Shown		
	Total Test	Knowledge Items	Attitude Items
80 - 99%	10	3	7
60 - 79%	4	4	0
40 - 59%	16	12	4
20 - 39%	19	11	8
0 - 19%	10	6	4
Total	59	36	23

The heterogeneity of the "after" MARK items to social work students is, as might be expected, less than the "before" version. Sixteen per cent of the items are significantly inter-correlated at the 5 per cent

level. At the one per cent level, ten per cent of the average inter-item correlations are significant.

The average inter-item correlation on the "after" MARK was also considerably higher than on the "before" form of the test.

The number of discriminating items on the "after" MARK was almost the same as on the "before" MARK, 47 or 80 per cent. The number which discriminated among all three thirds, however, was considerably larger and the separation between the top two-thirds and the bottom third is much more marked as shown in Table 15.

TABLE 15

ITEM DISCRIMINATION--"AFTER" MARK

	Discriminated Top Third From Middle Third and Middle Third From Lower Third	Discriminated Top Third From Other Two Thirds	Discriminated Lower Third From Other Two Thirds
Number of Items	26	4	17
Number of Attitude Items	6	2	9
Number of Knowledge Items	20	2	8

The I.D. Questionnaire and Interest Test

The reliability of the I.D. Questionnaire and Interest Test must rest on methodological grounds. The fact that these have been used for several years by the writer would tend to support the probability that particularly ambiguous items or those likely to be answered inconsistently would have been eliminated. The validity of these two instruments was

unknown at the time data was obtained. Approximately 20 per cent of the I.D. items correlated significantly with "after" MARK scores or course grades. Approximately the same proportion of interest items correlated with these variables.

Some indirect evidence of how well these two instruments discriminated among students might be deduced from the degree of correlation with another discriminating instrument like the "before" MARK. Because these two instruments, however, were aimed at measuring different variables than the "before" MARK, correlations with the MARK were smaller in number than would otherwise be the case. Stronger evidence for discrimination was thus shown by the fact that students were relatively well distributed on each of the questions in both of these questionnaires.

Forty per cent of the 32 items on the I.D. Questionnaire correlated with the "before" MARK; ten out of 64 interest items or 16 per cent correlated with this instrument.

The Special Test Questions

The special test questions were found to have a split-half reliability of .64, corrected for the number of items. Their validity, as measured by their correlation with more established measures of learning, like the research course grades was .34. Their validity as measured by their correlation with the "after" MARK was .33.[3] The discrimination and item difficulty of the special test questions were similar to that of the "before" MARK test, with item difficulty being somewhat greater.

---

[3]These correlations are corrected for attenuation of the criterion. Course grades were assumed to have reliability equal to the special test questions.

The Course Evaluation

The reliability and validity of the course evaluation, like the I.D. Questionnaire and the interest test, depends on its methodological soundness. Its use for several years eliminated a number of vague and dubious questions. Though correlation with other variables was precluded because students were not identified, the students' feeling of anonymity would be expected to increase validity of responses. Findings in previous years that student responses to this instrument in writing agreed with their oral evaluation of the course and the teacher would provide further support for its reliability and validity. The relatively large range of responses indicates satisfactory discrimination.

Research Course Grades

Unlike other measures mentioned above, the research course grades were not obtained by a single instrument. Further, they were computed by different teachers on the basis of varied information, and sometimes made use of what information was available with different degrees of precision.

The combining of students from the five schools had the problem of combining grades that were obtained on different bases and tended to move them toward the lowest common denominator of reliability and validity. This step, however, had at least the advantage of biasing the outcome in a conservative direction. Thus any prediction of grades from the "before" MARK scores would mean that the MARK could be expected to predict somewhat better if there were more reliable, valid, and discriminating criterion measures available of what had been learned.

Because some information on the reliability and validity of grades at the five schools was considered necessary to interpret the findings from this training plan, the writer made an attempt to categorize the five schools in this respect. In making these categorizations the writer assumed that grades based on objective tests were more likely to be valid than those based on short answer tests and that these in turn were probably more valid than those based on papers. It was further assumed that among schools using the same method of evaluating learning, those with the most experienced teachers, with teachers who had the most training in research, and whose field was closer to social work, would make the most valid evaluations of learning achieved. There is, of course, always the possibility that an experienced teacher can grade more reliably or validly on a term paper than an inexperienced teacher can construct and grade an objective test. On the basis of these assumptions, however, schools 2 and 5, and possibly school 1, are considered to have the most reliable and valid grades; schools 3 and 4 are considered to have grades which are least reliable and valid of the five.

CHAPTER VII

THE OUTCOME AND EVALUATION OF THE PROJECT

This chapter discusses the extent to which the two major objectives of this pilot training plan were reached. It includes an assessment of how well various levels of potential for learning research could be identified and what kind of teaching methods and content appear most likely to maximize each of the various potentials. It notes the relation between the students' evaluation of the course and measures of learning achieved.

Material showing the ability of the I.D. and interest data and the "before" MARK scores to identify and predict potentials for learning research is presented first. "After" MARK scores, course grades, and special test questions scores serve as criteria for these predictions. Analysis is then made of teaching methods to determine which best appears to maximize potential for all students, followed by analysis of those teaching methods that appear particularly useful for students at various levels of potential. In these latter two analyses, measures to evaluate the outcome of the pilot training plan are differences between the "before" and "after" MARK scores, special test questions scores, and student reports on course evaluations.

Some contamination of the study's outcome was considered possible because the MARK was originally designed at Tulane to measure potential for learning among the writer's students and to determine how this potential was maximized by the content of the beginning research course

as conceptualized by the writer at Tulane. This meant that correlations between the "before" and "after" MARK, between the "before" MARK and course grades, and changes in MARK scores, for example, might be higher at Tulane than at other schools. To test whether this situation significantly influenced results, all tests made on the five schools were repeated on the other four schools alone, with the Tulane students excluded from the sample. These tests did reveal the expected higher correlations and scores at Tulane. Nevertheless, all findings that were significant with the Tulane students included in the sample remained significant at the other four schools, with the Tulane students excluded. Thus any bias on the findings from the fact that the test was originally designed for Tulane and the method of teaching used there can be considered to be minimal. In the interest of brevity, findings on the four schools other than Tulane are not presented separately from those for Tulane and all findings in this chapter refer to all five schools as a group, unless specific results are labeled differently.

Identification of Levels of Student Potential For Learning Research

The writer considers that this pilot training plan was successful in showing that various levels of potentials for learning research could be identified and learning achieved by students at these levels predicted. Some evidence comes from both I.D. and interest data but the major identification comes from the "before" MARK scores.

Identification and Prediction by I.D. and Interest Data:

Though only a few of the I.D. and interest items identified levels of potential for learning research, the findings from those items which permit differential prediction are somewhat in line with findings from

previous studies.[1]

The I.D. items that predicted for all levels of students were those with face validity, and those found in these other studies to represent the students' social and cultural background and social class.

Three of the interest items that predicted show a liking for risks, billiards, fishing and playing poker.[2]

Table 16 below shows the correlations significant at the 5 per cent level between interest and I.D. items on the one hand, and the three criterion variables showing levels of learning achieved. Because these correlations are each based on only one item, they are considerably smaller than correlations between MARK scores and the same criteria for learning research.

---

[1]Robert H. Knapp, op.cit.

[2]David C. McClelland, "The Calculated Risk: An Aspect of Scientific Performance," Taylor and Barron, op.cit., pp. 184-217.

TABLE 16

CORRELATIONS OF INTEREST AND I.D. ITEMS WITH COURSE GRADES, "AFTER" MARK SCORES, AND SPECIAL TEST QUESTIONS SCORES

Item	Correlated With "After" MARK Scores	Course Grades	Special Test Questions
I.D. ITEMS			
Reaction to abstract reasoning	.14	.11	-
Previous course in research	.18	-	-
Previous course in research as an elective	.12	-	-
Previous course in logic	.14	-	-
Previous course in logic as an elective	.14	-	-
Education of father	.11	-	-
Father's occupation	-	.11	-
Mother's occupation	-	.11	-
INTERESTS			
Chemistry	-	.12	-
Baseball	.15	-	-
Coin collection	.11	.17	-
Billiards	.11	.17	-
Hiking	-	.11	-
Reading poetry	.17	-	.13
Square dancing	-	.17	.15
Flower arranging	-	-	.15
Fishing	-	.16	-
Auto mechanics	-	.12	-
Reading science fiction	-	.18	-
Playing poker	-	.18	-
Reading biographies	-	.16	.16
Reading science	-	.17	-
Church singing	-	-	.11

When interest items showing a liking for risks were combined to make a "risk score" this score correlated much more strongly with the criterion variables above than any of the interest items alone. This risk score was the number of interests that the students reported that appeared to involve risks. The maximum score was 6, shown by interests

in "studying the stock market," "betting on horse races," "playing poker," "playing bridge," "mountain climbing," and "fishing." Though this risk score was correlated only .18 with the final course grade it was correlated .36 with the "after" MARK score.

Identification of Potential Levels by "Before" MARK Scores:

In terms of their "after" MARK scores, the "before" MARK appeared to divide students into three groups. Slightly more than half of those in the top third on the "after" MARK scored in the top third on the "before" MARK; almost half of the middle third on the "after" MARK were in the middle third on the "before" test; almost two-thirds in the lowest third on the "after" form of the MARK came from the lowest third on the "before" form. Details are in Table 17 below. School to school variation in this pattern was small and not statistically significant.

TABLE 17

RELATION BETWEEN STUDENTS' RANK ON "BEFORE" OR "AFTER" MARK

Rank on "Before" MARK	Rank on "After" MARK*		
	Top Third	Middle Third	Lower Third
Top Third	51%	40%	9%
Middle Third	30%	45%	25%
Lowest Third	19%	15%	66%

*This kind of division seemed to produce more homogeneous groups than a more conventional division into quartiles.

Prediction of Learning by "Before" MARK Scores:

The potential of the "before" MARK to predict learning in research that was expected from the discrimination of the items, was borne out by its correlation with the three criteria for learning used in this project.

Prediction of "after" from "before scores. The ability of the "before" MARK to identify what kind of "after" MARK scores students will make is somewhat better than its ability to predict course grades, as shown by the correlation between "before" and "after" MARK scores in Tables 18 and 19 below.

The undergraduate grade point average predicts "after" scores considerably less well than "before" MARK scores.

TABLE 18

PREDICTIVE POWER OF "BEFORE" MARK SCORES WITH THE "AFTER" MARK SCORES AS A CRITERION*

Before Scores	Total Score After	Attitude Score After	Knowledge Score After
Total Scores	.69	.65	.57
Attitude Scores	.37	.35	.30
Knowledge Scores	.65	.61	.53
Undergraduate GPA	.28	.03	.37

*All correlations have been corrected for attenuation of the criterion according to the reliabilities given in Chapter VI.

Prediction of grades from "before" scores. Though the "before" MARK correlates less well with grades than with "after" scores, its predictive power with grades is comparable to or better than most tests

in use in other fields, as shown in Table 19 below.[3] In this case also the "before" MARK predicts significantly better, at the 5 per cent level, than the undergraduate grade point average.

TABLE 19

PREDICTIVE POWER OF "BEFORE" MARK SCORES WITH COURSE GRADES AS A CRITERION*

	Correlation with Course Grades
Total Score	.37
Attitude Score	.19
Knowledge Score	.43
Undergraduate GPA	.22

*These validities corrected for attenuation in the criterion only, represent the predictive power of the "before" MARK scores if grades had a reliability equal to that of total "before" MARK scores (.62). No correction for the reliability of the MARK has been made. Grades probably have a lower reliability than .62 so this estimate is on the conservative side and probably under-estimates the predictive power of the MARK. (For example, at the one school that used objective tests as the sole method of evaluating learning, the corrected correlations were .53, .32, and .47, respectively.)

Prediction of Special test questions scores from "before" scores. The ability of the "before" MARK to predict those students who did best on the special test questions is evidence of its ability to identify the group of students who might learn more than usually taught in the basic research course. These would be the students who could be expected to

[3]For example, the Miller Analogies, widely used in predicting success in Clinical Psychology, has been shown to have a correlation of .35 to .45 with course grades in psychology courses.

become researchers when they completed their Master's training.

"Before" MARK scores again predict special test questions scores considerably better than the undergraduate grade point average, as shown in Table 20 below.

TABLE 20

PREDICTIVE POWER OF THE "BEFORE" MARK WITH SPECIAL TEST QUESTIONS AS A CRITERION*

	Corrected for Attenuation of Criterion
Before Total	.31
Before Attitude	.22
Before Knowledge	.21
Undergraduate GPA	.08

*All correlations have been corrected according to reliabilities given in Chapter VI.

The Relation of the Teaching Variables to Various Levels of Potential

This part of the analysis deals with the outcome of the teaching in this pilot training plan. It assumes that the MARK is the best measure in this study that can be applied to all students, and that the achievement of potential is best measured by the change in MARK scores, particularly MARK knowledge scores, as explained below. The special test questions scores, on the other hand, are assumed to be the best measure of achievement of potential for the superior or outstanding group of students.

The course evaluation provided a self-assessment by students of

the learning they thought they had achieved, measures of their interest in the course, and their opinion on the usefulness of the content. In addition, it provided some data on how students viewed the teaching methods, and the amount of time they spent in studying.

This section, unlike the previous sections of this chapter, which have discussed the five schools as a whole, deals with differences among the five schools in the teaching variables and their relationship to these criterion measures. Because of the difficulty in quantifying data dealing with teaching methods, course content, and emphasis, this section has used dichotomous break-downs of the gains in learning made by students as compared to dichotomous break-downs on the teaching variables.

Similarity of the "Before" and "After" Sample:

The group who took the "before" and "after" MARK are considered to be representative of typical students in the five schools of social work under study despite the loss of 90 students between the two test administrations. The rationale for this conclusion is given below.

A test by chi square showed there was a significant difference between the group of 90 students who did not take the "after" MARK and the group of 263 students who took both the "before" and "after" MARK. This difference was significant at the 5 per cent level but not at the 1 per cent level. Comparison of the distributions of the two groups showed that almost all the difference in distribution was attributable to nine students in the group who did not take the "after" MARK. These nine students had scored extremely low on the "before" MARK also had very low undergraduate grade point averages and did not complete their graduate social work education. Thus, the group who took both the "before" and "after" MARK are considered representative of typical students at the five

schools involved.[4]

Differences Between "Before" and "After" MARK Scores:

Table 21 below shows that for four schools there was a significantly greater than chance (at the 5 per cent level or above) increase in MARK scores after the student had taken the research course.[5]

TABLE 21

INCREASE IN TOTAL MARK SCORES AFTER STUDENTS HAD TAKEN THE RESEARCH COURSE

School Number	Increase in Mean Score	Increase in Median Score	Level at Which Mean Increase was Significant	Level at Which Median Increase was Significant
1	3.4	3.0	.05	.05
2	3.1	3.5	.001	.001
3	2.9	4.5	.05	.01
4	0.9	1.0	-	-
5	9.1	9.0	.001	.001

Almost all of the increase in MARK scores was due to an increase in knowledge scores; attitude scores at four of the five schools showed a decrease, although this change was significant at only two of the schools, as shown in Table 22 below. (Changes in median scores were

[4]There were many reasons why the other 81 students did not take the "after" MARK. Some were ill or absent at the time it was administered. Some did not wish to complete it.

[5]Some of the lack of significant increase in school 4 may be due to the fact that only a third of the original students completed the "before" and "after" MARK under conditions similar to that at other schools.

quite similar to changes in mean scores and in the interest of simplicity are omitted.)

TABLE 22

INCREASE IN MEAN ATTITUDE AND KNOWLEDGE SCORES AFTER STUDENTS HAD TAKEN THE RESEARCH COURSE*

School	Increase in Attitude	Level of Significance	Increase in Knowledge	Level of Significance
1	-0.6	.10	3.0	.001
2	-0.5	.10	3.5	.001
3	-1.7	.01	4.6	.001
4	-1.7	.01	2.6	.10
5	1.0	.01	8.0	.001

*Minuses before numbers indicate a decrease in mean scores.

The decrease in attitude scores, or the failure of attitude scores to increase, may be related to the lack of reliability of the attitude part of the MARK or to other factors. Some of these other factors may be the students' attempt to show attitudes they thought were desired, or the fact that some schoolsstressed different attitudes than the ones used in scoring the MARK.

Frequency distributions for each of the five schools showed similar changes in range of scores. To preclude identification of schools by sample size, frequency distributions are not listed here.

The "before" MARK total scores ranged from 12 to 41, out of a possible score of 59. Twenty-two scores were below 20 and only three were 40 or above. Out of a possible score of 59 on the "after" MARK, the range was 16 to 51; only five were below 20, and 30 students scored 40

or above.

Relation of increase in MARK score to students' rank on "before" MARK. A disturbing inverse relationship was noted between the amount learned, as shown by differences between "before" and "after" MARK knowledge scores and the rank of students on the "before" MARK. If the goal of the research course is to maximize learning for the top third as well as for the other two thirds of the students, rather than bringing each student up to some minimum standard, this situation presents a particular problem for the profession.

Students in the top third have the lowest gain in knowledge, as shown by MARK scores; students in the bottom third have the most gain; while students in the mid-third were between these two groups. See Table 23 below.

TABLE 23

INCREASE IN MARK KNOWLEDGE SCORES BY STUDENTS' RANK

Students' Rank	Mean Knowledge Score "Before" MARK	Mean Knowledge Score "After" MARK	Increase
Top third of students	20.1	23.3	3.2
Middle third of students	15.9	21.0	5.1
Lowest third of students	11.6	19.0	6.4

When comparisons between sections are made at the two schools that assigned students to sections by "before" MARK scores, schools 2 and 5, the same inverse relationship between rank on the "before" MARK and learning achieved is found, even though one school showed a considerably higher gain than the other. See Table 24.

TABLE 24

INCREASE IN MARK SCORES--STUDENTS SECTIONED BY RANK

Sections	Mean "Before" MARK	Mean "After" MARK	Increase
High section-School 2	17.8	19.6	1.8
Low section-School 2	14.3	17.1	2.8
High section-School 5	18.9	24.9	6.0
Low section-School 5	11.4	21.0	9.6

The failure of students who scored higher on the "before" MARK to gain as much as their colleagues who scored lower on the "before" MARK does not appear to be due to the inability of the MARK to measure learning achieved. Even those students who scored highest on the "after" MARK, missed from 10 to 15 of the 59 questions so if they had additional knowledge, there was opportunity for them to indicate this by answering these questions correctly.

Instead, the relatively small gain of those students scoring high on the "before" MARK appears partly related to the level of content in the research course.

The writer believes that the relatively high anxiety students express about learning research and the fear of some faculty for them in the course both tend to reduce the level at which the research course is taught. This climate, plus the boredom at hearing already known material, may well result in only a little learning among students with potential for much more gain.

Relationship Between Changes in MARK Scores
and Teaching Variables at the Five Schools:

Changes in median knowledge scores were used as the major variable in analyzing the outcome of this pilot training plan. Median scores were selected because they were less influenced by unusually outstanding or limited students. Mention is made, however, of findings that are supported or not supported by changes in mean scores.

Conclusions suffer from all of the problems of this project design which tried to exploit individual differences among schools and sections as experimental variables. The writer recognizes that conclusions are made in considerable ignorance of interaction effects among the various variables, and of unknown influences from uncontrolled variables. Nevertheless, they appear to offer a beginning point for further pilot training plans or research studies.

Median knowledge scores increased most at schools 2 and 5 and least at schools 1,3, and 4. Mean scores increased most at schools 3 and 5 and least at schools 1,2, and 4. Thus school 5 had the greatest increase in both mean and median scores, while schools 1 and 4 had the least increase in these two scores.

The material below first presents findings for each school as compared to each other school. These are followed by comparisons of sections within schools.

The method of teaching seemed to make only slight difference in student learning, but schools with most didactic teaching made slightly higher scores. The most significant finding about teaching methods was that exercises at home appeared to have little influence on knowledge as measured by the MARK.

Schools who used objective tests had more increase in knowledge scores than those who used papers as a bases for grades.

Listening to lectures appeared to be the best learning experience for obtaining an increase in knowledge scores.

Emphasis on content that covers techniques and instrumentation appeared to make for most learning as measured by the MARK.

Teacher characteristics made very little difference in MARK scores. Most increases in scores were at schools with teachers who had moderate experience.

More hours of credit and more hours of additional electives open to a student in research tended to be associated with increased student learning; courses early in the students' educational experience were associated with less changes in MARK scores than those which occurred later. The often controversial factor of class size appeared to make very little difference in gains in knowledge as measured by the MARK; if there was any difference, it slightly favored the larger classes.

Details on these findings follow:

Teaching Methods. With the exception of exercises in class, didactic teaching appeared to have a slight advantage over discussion methods of presenting material. Exercises in class appeared to have considerably more association with increase in MARK scores than exercises at home. Least influence on changes in MARK scores came from class demonstrations by the teacher. These findings were supported by changes in both median and mean scores.

Learning Experiences. Findings on learning experiences fitted in with the findings about teaching methods. More changes in median and mean MARK scores occurred at those schools where students' learning

experiences were largely listening to lectures in comparison to those which had more discussion. The relatively minor pertinence of reading research studies as a learning experience to changes in MARK scores may be due to the failure of the MARK to identify this kind of learning; the analysis of studies on the other hand, appeared to be associated with upward changes in both mean and median scores. Schools that emphasized analysis of studies had more change in these scores than those which did not.

Evaluating Learning Achieved. Those schools with the most objective measures of evaluating learning made most gain in median MARK scores; those with least objective measures of evaluating learning made least gains in median scores. Changes in mean scores were less consistently associated with these findings, but followed this pattern.

Content. All schools that emphasized research techniques and instrumentation had larger increases in median scores than schools which do not. This finding was upheld to a partial extent with regard to mean scores. Least pertinent content to increasing MARK scores was statistics and phases in the scientific method; schools that emphasized these most had less gain in both median and mean scores than schools which emphasized them least. Content highlighting the findings of research studies was only slightly associated with changes in MARK scores.

Teacher Characteristics. Gains in MARK scores were only slightly related to teacher characteristics. Less knowledge as measured by the MARK was gained in schools with teachers who had most and least training and experience than where teachers had moderate training and experience.

Miscellaneous Variables. Four of the miscellaneous variables had relationships to changes in MARK scores. Consistently, in those schools which had the most hours of credit, gain in MARK scores was highest.

Where the number of hours of credit for research courses was smallest, the change in MARK scores was least. At schools providing most remedial work and the greatest number of elective hours of research on completion of the basic course, there was a greater gain in median scores than in schools offering less. On the other hand, where research was taught early in the students' educational experience, there was less gain in MARK scores than in the schools where it was taught later.

Relationship Between Changes in MARK Scores and Teaching Variables in the Fourteen Sections:

Except when differences were deliberately introduced for experimental purposes, all participating schools made considerable effort to control differences in course content, emphasis, and requirements for papers and examinations, in order to keep inter-section differences within each school to a minimum.

Earlier mention has been made that the relatively small increase in knowledge by the top third of students on the "before" MARK (as measured by differences between "before" and "after" knowledge scores) appeared to be a function of the content of the conventional research course.[6]

What the additional content should be for students with some background in research is not yet clear, but it appears likely that these students can and will learn more about how to conceptualize, the relation between theoretical and empirical concepts, and techniques and instrumentation in research than they will about statistics.

---

[6]This finding is similar to that of L.D. Jaffe, in "Building an Undergraduate Social Work Research Sequence," Journal of Jewish Communal Services, Vol. 42, No. 1, 1965, pp. 99-108.

At school 2, for example, Section A had considerable statistical content added as an additional requirement. Despite the fact that about 20 per cent of the MARK measures knowledge of statistics, this section gained less in knowledge as shown by MARK scores than any other section at any other school. At school 5, Section A had considerable additional content added about theoretical and empirical conceptualization and research techniques and instrumentation that was considered to be at an advanced level. These students made a greater gain on the MARK in both mean and median scores than any other section except Section C at school 5. Some of the students in Section C had considerable remedial help from tutors and this assistance may have been responsible for this relatively greater increase in scores.

Influences of teacher differences are most noticeable when sections are examined. Sections that had teachers with social work background generally gained more than sections with teachers from some other field. Within a school, sections that had more lecture appeared to gain more than those with more discussion.

Relation Between Change in MARK Scores and Special Test Questions:

The lack of relationship between special test questions scores and changes in MARK scores would indicate that the STQ's measure something other than that measured by the MARK. Further support for this belief is shown by the findings that those students who scored highest on the "before" MARK (and thus showed most potential for learning research) had higher scores on the STQ's than those students who scored lowest on the "before" MARK.

Special test questions scores were obtained at only four schools. School 4 did not give this test.

There was little relation between the median change in MARK scores and the STQ scores. School 5 was high on both the change in MARK scores and the STQ scores; school 1 was low on both scores. School 3, however, was low on the change in MARK scores and high on the STQ scores; school 2 was high on the change in MARK scores and low on the STQ scores.

Students who scored in the highest third on the "before" MARK made special test questions scores significantly higher (at the 5 per cent level) than those who scored in the lowest third on the "before" MARK. The gain in students' scores on the "after" MARK was negatively associated with their STQ scores. That is, those students who made the most gain in "after" MARK scores had the lowest STQ scores or those who made the least gain had the highest STQ scores.

Relation of Special Test Questions Scores and Teaching Variables by Schools:

Only a few of the teaching variables were related to the special test questions scores.

The schools with most exercises in class '3 and 5) and with teachers with more social work background (1 and 5) ranked higher on the special test questions than those with least class exercises or teachers from outside social work. Schools with fewest lectures and those with most emphasis on reading research studies and statistics, (2 and 4) made lower special test questions scores than the others.

Section to section variations in special test questions scores were similar to those among schools. Section A at school 5 in which additional advanced level content was added, scored highest on the STQ's. Students in Section A at school 2 where content on statistics was added, scored lowest of all sections on STQ's.

Students' Responses to the Course Evaluation

Students' Assessment of the Knowledge They Gained:

After completing their research course work, only three out of ten students believed they were ready for an advanced course in research; one-half were uncertain and the remainder did not feel ready. About half of the students felt ready to participate in research as a member of a team, and three out of ten thought they could carry on a research project with some consultation.

Usefulness of Students' Learning to Their Practice:

Almost eight out of ten students considered the research course less useful to practice than other courses, but six out of ten reported they received more help from this course than they had expected. Six out of ten students believed that research should be a required rather than an elective course. Nine out of ten students were of the opinion that research should come in the first year of the Master's program and should be of at least a full semester's duration.

Students' Interest in the Course:

Responses on three items measuring interests were almost identical to those on items measuring usefulness of learning to practice.

Students' Reaction to Teaching Methods:

Only about one-half of the students felt as free to raise questions in class as they would have liked, but seven out of ten found teachers' answers to questions as helpful as could be expected. With regard to specific teaching methods such as lectures and discussions, students were generally satisfied with what they were getting. Half of

the students, however, wanted more demonstrations by the teacher and exercises in class and fewer exercises at home. About the same proportion indicated a preference for a seminar or class smaller in size than theirs. Class size was reported as the chief obstacle to student learning. As might be expected, this reaction was much more prevalent in schools with larger classes.

Seven out of ten students reported they spent less than three hours in preparation for class. The average preparation for examinations was an additional eight to twelve hours of study.

Students' Characterization of Course Content:

Students' ratings of the course content with regard to its usefulness to their practice, difficulty in learning, liking by the student for the content, material requiring most study time, and the quality of the teaching showed much variation.[7]

Content about parameter-statistic relationships was characterized most frequently as most difficult, least liked, taught least well, and requiring the most time for study. No other content was characterized with any similar outstanding frequency. Least difficult content, requiring fewest hours of study was sampling. Most useful in practice was reading and analyzing research reports. Teachers were said to

---

[7]Topics from which students could make a choice in these characteristics included the following: scientific approach to knowledge, the cycle of scientific inquiry, question and hypothesis formulation, standards for observations, methods for making observations, connecting questions and hypotheses to observations, qualitative classifications, rating scales, sampling, research design, empirical conceptualization, measures of central tendency and variation, relationships between variables, parameter statistic relationships (significance tests), theoretical conceptualization, connecting theories and observations, schedule making, research interviewing, writing research proposals and research consultation, reading and analyzing research reports.

teach question and hypothesis formulation best. No single content area was an outstanding favorite. Writing research proposals and materials on research consultation was liked slightly more than other topics.

While it is possible that some students at schools other than the writer's interpreted the meaning of categories of content given them differently than students at the writer's school, school to school comparisons show considerable similarity, as shown by Table 25 below.

TABLE 25

SCHOOL TO SCHOOL COMPARISONS ON STUDENT CHARACTERIZATION OF COURSE CONTENT

Most difficult--parameter statistic relationships--Schools 2,3,4, and 5; scientific approach to knowledge--School 1.
Least difficult--measures of central tendency--Schools 1,2, and 3; sampling, Schools 2,4; the cycle of scientific inquiry, School 5.
Liked best--no consensus--scientific approach to knowledge, School 1; writing research proposals at School 2; reading research reports at School 3; no consensus at School 4; scientific approach to knowledge at School 5.
Liked least-- parameter statistic relationships--all schools.
Most useful--reading and analyzing research reports--Schools 1,2,4, and 5; connecting theories and observations, School 3.
Least useful--parameter statistic relationships--all schools.
Most time spent studying--no consensus--scientific approach to knowledge, School 1; measures of central tendency, School 2; research design, School 3; reading reports, School 4; research interviewing, School 5.
Least time studying--no one topic received a majority vote at any school.
Taught best--no consensus--scientific inquiry, School 1; design, School 2; question and hypothesis formulation, School 3; cycle of scientific inquiry, School 4; writing research proposals, School 5.
Taught least well-- parameter statistic relationships, Schools 2,4, and 5; scientific approach to knowledge, School 1; empirical conceptualization, School 3.

Relationship Between Change in MARK Scores and Students' Evaluation of Course:

There was almost no relationship between change in MARK scores by schools and the way schools were divided on the items used in the students' evaluation of the course. For two items where a relationship was present, one was negative and the other positive. There was a slight negative relationship between the students' report of interest in research, for example, and the change in MARK scores, with less interested students showing the most change. On the other hand, there was a slight positive relationship between the changes in MARK scores and students' assessments of the usefulness of the course.

Because course evaluations did not identify students by sections, no analysis of the relationship between findings from the course evaluation and sections was possible.

Relationship Between Course Grades and Course Evaluations:

Since each student recorded his research course grade on the evaluation, when it was available, it was possible to classify course evaluations with regard to whether they were by superior, average, and below average students. For this analysis, students with "A+," "A" and "A-" (38%) were put in one group; students with "B+" and "B" (47%) in another; and the "B-," "C" and "F" students (15%) made up a third group. In the material below, these groups are referred to as "highest," "middle" and "lowest" groups. Grouping students in this way by grades was considered to minimize problems from the reliability and validity of grades.

Student Assessment of the Knowledge They Gained. The lowest group had most conflict about their ability. They felt considerably less ready

to go on to an advanced course than the other two groups, but expressed the same degree of confidence as the other groups that they were ready to participate in a research project.

Usefulness to Practice and Interest in Research. A larger proportion of the middle and highest group believed research should be required, but only the highest group wanted two semesters of research required instead of one. The middle group found the course more interesting than the other groups; a larger proportion of the highest group, than in the other two, found it more interesting than they had anticipated.

Reaction to Teaching Methods. The highest group felt most freedom to participate, but still felt most restricted by class size. They wanted less lecturing, fewer direct answers by the teachers to questions from students, more class discussion and smaller classes. They preferred objective tests.

The middle group found the teachers' answers to questions most helpful, wanted more demonstrations in class by the teacher, smaller classes and a slower pace to the course. They preferred that grades be based on both papers and tests.

The lowest group felt most restricted by the pace of the course and wanted more exercises at home.

Though students in the lowest group reported that they spent most time in preparation for class, there was little difference between the middle and highest groups on this variable. Time spent in preparing for examinations was inversely related to the rank of the three groups.

Students' Characterization of Course Content. The three groups differed considerably in how they characterized content as shown in

Table 26 below.

TABLE 26

CHARACTERIZATION OF COURSE CONTENT BY STUDENT LEVEL

	Highest	Middle	Lowest
Most difficult	Parameter-statistic relationships	Parameter-statistic relationships	Parameter-statistic relationships
Least difficult	Measures of central tendency	Sampling	Sampling
Liked best	Scientific approach to knowledge	Methods for making observations	Methods for making observations
Liked least	Parameter-statistic relationships	Parameter-statistic relationships	Parameter-statistic relationships
Most useful	Connecting theories and observations	Reading and analyzing research reports	Reading and analyzing research reports
Least useful	Parameter-statistic relationships	Parameter-statistic relationships	Parameter-statistic relationships
Most time in studying	Parameter-statistic relationships and research designs	Connecting theories and observations and reading reports	Scientific approach to knowledge
Least time in studying	Sampling schedule making	Research interviewing rating scales	Sampling the cycle of scientific inquiry
Taught best	Cycle of scientific inquiry	Design and question formulation	Scientific approach to knowledge
Taught least well	Parameter-statistic relationships	Parameter-statistic relationships	Connecting theories and observations

Relationship of Students' Evaluation of Learning and Teaching Variables by Schools:

Student Assessment of Knowledge They Gained. Didactic teaching appears to give students less confidence but a more realistic assessment of their ability and better preparation for other work. At schools with didactic teaching, students were more likely to report they were ready for an advanced research course than those with discussion teaching methods. Yet, students taught didactically, felt confident only to engage in research as a member of a team and not to do a complete study even with consultative help. The method of evaluating students' learning had no relation to students' assessment of their knowledge. Content about the scientific method seemed most likely to increase students' confidence in their knowledge and the reality of this confidence.

The teacher characteristics of long experience, most training, full-time work, and social work background appeared associated with most realistic student assessment of their own research knowledge. Of the miscellaneous variables, only class size seemed associated with students' assessment of their knowledge gained, with larger classes having the most realistic assessment.

Usefulness to Practice and Interest in Research. Schools with most discussion methods of teaching, course content from studies, less experienced teachers, and small classes reported most usefulness for content and most interest in it.

Reaction to Teaching Methods. Students found teachers' answers more helpful in the class with most didactic teaching, where content emphasized phases in the scientific method, and where the teachers had more experience.

Students who were taught didactically spent more time in preparing for their courses than students in discussion classes. The same was true of students who were given objective examinations versus those who were asked to write papers. The only kind of content related to time spent in preparation was that relating to scientific method; students at schools with more emphasis on scientific method spent more time in preparation than at schools emphasizing other content. Students consistently spent more time in preparation for classes taught by teachers who were more experienced and had more training than with teachers of less experience and training. The only miscellaneous variables associated with time in preparation were class size and remedial work. Students in larger classes as well as those who had most remedial work appeared to spend most time in preparation.

Students' Characterization of Course Content. Only one relationship between student characterization of course content and teaching variables at the various schools was found. While in all schools parameter-statistics relationships was the most difficult irregardless of emphasis, students at those schools which emphasized statistics and quantitative measures found measures of central tendency and variability least difficult while others did not.

CHAPTER VIII

DISCUSSION OF FINDINGS AND SUGGESTIONS FOR THE FUTURE

The evaluation of this pilot training plan indicates that identification of various levels of potential for learning research among social work students can be made. This identification has somewhat more reliability and validity than the determination of those teaching variables that maximize learning for students at each of the various levels. Suggestions for current teaching, as well as for future pilot training plans, come from both findings.

This project has provided evidence of the ability of the MARK test to divide students into at least three levels in terms of their potential for learning social work research. At the same time, it indicated that the MARK might discriminate better if some attitude items were changed or eliminated, and if more knowledge items were added. Further work on this instrument to implement these suggestions is planned and the companion monograph to be completed approximately July, 1968, should show how this improved form of the MARK can separate students more discriminately according to their potential for learning research. The second monograph will also attempt to explain why the MARK predicts and to develop some beginning theory on what makes for success in various aspects of social work research.

Three Kinds of Student Potential for Learning Research

The writer has noted three rather distinct kinds of students who appear to correspond generally to the three levels of potential for

learning research that have been identified by the MARK.

In the first group, are those students who scored highest on the MARK, scored somewhat better than average on the attitude items, and whose performance in the research class indicated a strong interest in the "doing" aspects of research.

In the second group, were those who scored almost as high on the MARK and whose potential for learning research was almost as good as the first group, but whose interest in actually engaging in research activities was considerably less. Although this second group had less interest in actually engaging in research themselves, they recognized the importance of research based knowledge to forward professional goals, and they thus had more than a passive interest in supporting research when it was carried out by other members of the profession.

The third group had considerably less potential for learning research and almost no interest in doing it. This group, however, showed some interest in using the products of research, and a potential for learning enough research to use its products effectively in their professional practice.

The writer has designated these three groups as "doers," "supporters," and "users" on the basis of their interests and potentials.

Members of the "doer" group are seen as having a potential, with some consultative help, for engaging in social work research after completing a Master of Social Work degree.

The "supporter" group are those who could be expected to support and forward research activities carried out by others, although they will probably not engage in it themselves. The students in this group appear to be those who will quickly find their way into supervisory and

administrative positions after graduation. In these latter positions, they will be in a position to support and forward research carried out by others.

The "user" group represents the third level of the research potential identified by the MARK. These are the students with limited ability to learn research but who appear able to use its findings in their practice. Use of research findings is considered an inter-actional process, so these persons are seen as more than "consumers" of research. They would be expected to indicate how research findings met their practice needs and where findings needed reevaluation or extension. From this group one could expect suggestions that would guide the work of the "doers."

Among these three kinds of students, the most important finding of this study is the apparent limitations of current teaching methods and/or content to maximize the learning for the "doer" group. Evidence for this was the relatively small change between "before" and "after" MARK scores for the "doer" group, as compared with changes in these scores found in the "supporter" and "user" groups, and in the evaluative comments of all three groups.

Students in the "doer" group not only gained less knowledge than either of the other two groups but wanted more research courses.[1] Students in the "doer" group were dissatisfied with the pace of the

---

[1]This finding is supported by a follow-up study of 83 students carried on at Tulane in 1966-67, data for which became available during the writing of this report. These 83 students were placed in sections according to their "before" MARK scores. The mean increase in MARK scores in these four sections corresponding to the four quartiles on the "before" MARK was as follows: 1st quartile 3.8; 2nd quartile 8.1; 3rd quartile 11.2; and 4th quartile 13.0.

course, and wanted a faster pace and more content. They had less interest in the course and found it less useful than other groups.

The finding that the high scoring students advanced least supports the writer's previous undocumented supposition that the potential of those students who could learn most was not being realized. Concern about this point was one of the motivating factors for this project.

This inverse relationship between the amount of knowledge held at the beginning of the course and the gain in knowledge during the course appears to be associated with lack of sufficient course content to meet these students' needs; or if the content is in the course, this lesser learning appears related to lack of stimulation of student learners.

In 1962 through 1964, at Tulane, the writer tried teaching research to the students with highest MARK scores in sections where they were mixed with other students; in 1965 and 1966 these high scoring students were placed in a section by themselves, and additional content was introduced into this section. As measured by change in MARK scores, the learning of these high scoring students was less than other students in both instances.

The writer now believes that some different approach is necessary. He suggests these high scoring students on the MARK have the alternative of taking something other than a beginning course in research, either an intermediate research course if this fits their interests or some elective other than research if they are less interested in research than some other area.

Comments of those students who scored high on the "before" MARK indicate that at least some of the failure of top-scoring students to learn more is related to their lack of stimulation. This situation

occurred even when additional material was added for these students by the writer in 1965 and 1966, because some of them insisted that the course also include a review of beginning content. If high scoring students were permitted to have a choice of electing an intermediary course of research or some other kind of course, a more homogeneous group for each kind of course could be an outcome, with more stimulation possible, and consequently greater learning.

The special test questions scores show that high scoring students who can be considered better prepared can learn additional content even if it is part of a beginning course. If an intermediary course was built on what they already knew, was limited to students who elected it, and thus exploited their interests, greater learning could be expected. On the assumption that these students know what kind of teaching is best for them, it is suggested that any such additional course be taught in a discussion manner with objective examinations used to assess learning.

Further study is needed of the influence of specific content emphasis on student learning. In this project not enough detail was available on specific content taught at the various schools, and not enough was known about which items on the MARK measured various content aspects to study this relationship. At the same time further experimentation can be carried out with special courses and content planned for the three groups of "doers," "supporters," and "users."

A Suggested Curriculum to Maximize Student Potential

Such a suggested experimental curriculum is outlined below. Further study can be made, both of the influence of specific current content emphasis, and this kind of specialized experimental teaching on

student learning. The MARK and other instruments developed in this study can be used as measuring devices of what knowledge students gain.

The writer visualizes a research sequence that will provide two semesters of three hours each of classroom work in research and with two separate projects or experiences of at least one semester each. The first semester of classroom work would be required for all students except those in the top third on the MARK. These would be exempted, and for them the second semester would be elective. The first project would be required for all students; the second project would be elective.

The first semester would have as its objective the preparation of all three kinds of students for the various professional activities of doing, supporting, and using research. Most emphasis would be given in this beginning course to preparing the students to be "users." This stress would be based on the assumption that all students would be expected to be "users" to some extent. Content relating to doing, supporting, and using research would be included, but content relating to using would be emphasized most. Teaching methods most appropriate for each kind of content would be used for the particular content injected. Thus, the entire semester would not be lecture, or discussion, or exercises, etc., but presentation and learning experiences would relate to each of these three kinds of content and objectives.

Student interest in research and integration of research learning with the rest of the curriculum would be forwarded by having a practicum or research experience beginning either at the same time or shortly after the course began. This research experience would be closely related to the students' field instruction so that the topic of research would emerge from the students' field placement experience. Field instructors and

research instructors could jointly work on this project. Field instructors could use the research experience to show how helping clients can be facilitated by research based knowledge; research teachers can use it to apply some of the theoretical aspects of research covered in the research class.

Integration of research into the field experience could be expected to provide a number of benefits to students. In the field, the student sees situations about which he needs more knowledge or aspects of which he wants to change. Engaging in a research project related to these situations should help the student to recognize the importance of having knowledge before decisions are made, or to see the usefulness of knowledge in affecting change.

Students at the other two levels, supporters and doers, would have additional learning experiences available as electives. Both groups would have an opportunity to elect a second research classroom course. Content of this course would cover such topics as how to obtain funds for research (and would include exercises in writing grant requests), budgeting and administrative activity related to research, the role of research in the administrative structure of the agency, the use of research consultation, etc.

Doers could elect this course plus a second practicum, or substitute field instruction in research in their second year for field instruction in one of the other practice methods.[2]

---

[2]Tulane has begun to develop such a sequence on an experimental basis. Informal evaluation of student reaction to such a plan is quite positive. Further work on this plan is now going on and interested persons can obtain more details from the writer.

Suggestions for Future Training Plans

With this proposed curriculum as a basis, the writer suggests further pilot training plans with course content, teaching methods, learning experiences, etc., based on what has been learned so far in this project.

A similar number of schools would appear appropriate for such a training plan; the size of the sample for this project was large enough for accurate statistical calculations but small enough to be manageable. More preparation for the training plan, such as more prior discussion between and among teachers about specific content and earlier planning by schools for experimental teaching could increase the validity of findings. The design could also include better control of certain variables so that section to section evaluation would not be confounded by undesirable combinations of experimental variables. For example, comparison between didactic and discussion methods of teaching could be made, with the size of the class held constant, or with a design that permitted variation in both these variables. Using both objective examination and papers for the same class would permit comparisons of these methods of evaluating learning with learning as measured by the MARK scores.

There is evidence from the analysis of the outcome of this training plan that additional content concerning conceptualization, the relationship between theoretical and empirical models, techniques, and methods of instrumentation, obtaining funding for research, and the use of research consultation can all be introduced for all three levels. Because material on statistics seems to inhibit learning in the course some other method of introducing it should be attempted. Tulane has

tried to limit the research classroom course to discussions of the meaning of statistics and introduce the computations of statistics into the project at the time when a student feels the need for such computations to test his hypotheses. This change, however, has not been in effect long enough to permit evaluation of outcomes.

It is likely in carrying out such a training plan, that some students with weak backgrounds in research and scientific method, and with poor undergraduate records will have even more difficulty in the course than they do now. In that event, remedial help should be provided, preferably by tutors recruited from second year students or students in an advanced program. Values of this tutoring plan would be threefold: assistance for the beginning students in difficulties; teaching experience for the students providing the help; and for the teacher, freedom to aim instruction at a loftier target than the "lowest common denominator" of the class.

Evaluation of this proposed training plan would be carried out in a manner similar to that used in this study, with the modified instruments that will shortly be available. It is expected that further item and content analysis will result in scoring patterns particularly characteristic of the three groups, the "doers," the "supporters," and "users." These additional scoring patterns will provide further criteria for evaluation of the outcome of teaching. Further changes in the course evaluation form can fit it to changes in teaching methods and content.